注：本文为胡传鹏博士的博士毕业论文，推荐引用格式为：

胡传鹏. (2017). 道德自我在认知加工中的优势. (博士), 清华大学, 北京. 获取自 http://www.chinaxiv.org/abs/201909.00137

Or,

Hu, C.-P. (2017). The Priority of Moral Self in Cognitive Process. (PhD), Tsinghua University, Beijing. Retrieved from http://www.chinaxiv.org/abs/201909.00137

如有任何疑问，请联系：胡传鹏，Email: hcp4715@hotmail.com

道德自我在认知加工中的优势

（申请清华大学法学博士学位论文）

培 养 单 位： 社会科学学院

学　　　科： 社会学

研　究　生： 胡 传 鹏

指 导 教 师： 彭 凯 平　教　授

二〇一七年六月

道德自我在认知加工中的优势

胡 传 鹏

The Priority of Moral Self in Cognitive Process

Dissertation Submitted to

Tsinghua University

in partial fulfillment of the requirement

for the degree of

Doctor of Philosophy

in

Sociology

by

Hu Chuan-Peng

Dissertation Supervisor： Professor Peng Kaiping

June, 2017

关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：

清华大学拥有在著作权法规定范围内学位论文的使用权，其中包括：（1）已获学位的研究生必须按学校规定提交学位论文，学校可以采用影印、缩印或其他复制手段保存研究生上交的学位论文；（2）为教学和科研目的，学校可以将公开的学位论文作为资料在图书馆、资料室等场所供校内师生阅读，或在校园网上供校内师生浏览部分内容；（3）根据《中华人民共和国学位条例暂行实施办法》，向国家图书馆报送可以公开的学位论文。

本人保证遵守上述规定。

（保密的论文在解密后遵守此规定）

作者签名：＿＿＿＿＿＿＿＿　　导师签名：＿＿＿＿＿＿＿＿

日　　期：＿＿＿＿＿＿＿＿　　日　　期：＿＿＿＿＿＿＿＿

摘 要

道德自我指个体对自己在道德方面的评估。作为自我概念的核心部分，道德自我对人们的道德判断和道德行为均有重要影响。最近的研究表明，道德自我相比自我的其他方面具有优势，例如自我增强效应（self-enhancement effect）在道德方面比其他方面更强，即人们普遍认为自己在许多方面处于平均水平之上，而在道德方面这种优越感最强。但是，道德自我的优势是否也存在于认知加工中仍是未知的问题。本研究结合认知神经科学的方法，以社会联结学习为研究范式，系统地考查了道德自我在知觉中的加工优势，提出道德自我是指导个体行为的内在参照标准，本论文从行为学（第二~四章）和神经生物学（第五章）水平为这一理论提供实证证据。

首先，通过三个控制研究（第二章：研究 1-3），验证了道德品质信息在知觉加工中的优势，同时排除了熟悉性和文字启动等可能的混淆变量，为道德信息能够对知觉起到自上而下的影响提供证据。如果道德自我是个体行为与认知加工的内在参照标准，则研究 1-3 的结果也受到这一标准的调控。因此，我们预期，当把自我与他人作为外显变量时，道德信息对知觉的调控效应主要反映在道德自我这一条件上。为了验证这一假设（第三章），研究 4 比较了道德意义上积极、中性和消极的自我以及他人在知觉匹配中加工的差异，发现道德品质信息对知觉的调节作用主要发生在与自我相关的条件。为进一步将道德信息与自我相关的交互作用，分别将道德（研究 5）或者自我（研究 6）作为与任务无关的刺激，检验道德与自我对知觉的内隐调节作用，结果发现道德与自我能够对知觉产生调节作用。为了讨论道德自我在知觉加工优势效应的稳定性，我们采用元分析进一步表明（研究 7），道德自我信息相对于其他条件人加工优势在反应时上的效应量大约为 Cohen’s d = 0.5 左右，表现出中等的效应量。

为了深入理解道德自我的知觉机制，本研究采用模型分析的思路分解认知过程（第四章），使用扩散模型和指数高斯模型对研究 4～6 的数据分析，研究 8 发现道德自我对知觉决策的影响主要是通过加快道德自我条件下的信息累积速率，而不是改变被试的决策偏见。为考察道德自我的知觉加工中的跨任务稳定性，在研究 9 结合了知觉匹配任务与知觉分类任务，且在知觉分类任务中设置了多种标准，结果发现，道德自我在不同的任务中均表现出稳定的优势。

论文的第二部分，进一步研究了道德自我的神经机制。一方面，使用大样本的神经成像数据，探索了自我道德评估得分与大脑结构的关系（第五章）。通过基于体素的形态学分析（研究 10），我们未能发现，道德自我评分与大脑灰质体积之间的相关。为了对参与道德自我任务的脑区进行评估，我们使用神经成像元分析（研究 11）比较了道德认知与自我参照的神经成像研究，发现两者共同激活了背内侧前额叶和眶额皮层，为进一步功能成像研究提供了良好的基础。

本课题使用社会认知神经科学的方法，从功能、算法和硬件三个层次研究了道德自我概念，将社会心理学与认知心理学和计算模型结合起来。本研究的一系列实验表明，道德自我在知觉中具有稳定的优势，这种优势表现为积极的道德自我信息在知觉中能够更快地进行信息的累积，其神经基础可能是背内侧前额叶和眶额叶皮层。道德自我在认知加工中的优势，与社会心理学中自我增强的效应一起，共同反应了道德自我概念对于个体的重要性，为今后进一步改变道德行为提供了基础。

关键词：道德自我；知觉决策；计算模型；神经成像；

Abstract

Moral self, which is a core part of self-concept, is the overall self-evaluation in the moral domain. Previous studies in social psychology has shown that moral self is closely related to individual's moral behaviors. Interestingly, recent studies found that moral self showed advantages over other aspects of self. For example, it was reported that self-enhancement effect in moral domain is stronger than that in other domains. However, it is still unknown whether moral self related information were processed preferentially during cognitive process. To answer this question, the current study investigated the enhancement effect of moral-self in perceptual processing, by adopting a cognitive neuroscience approach. Based on our behavioral (Chapter 2 ~ 4) and neural (Chapter 5) results, we proposed that moral self served as internal reference for individual's cognitive processing.

Firstly, the current study confirmed that the moral related information is processed preferentially during perceptual processing by three experiments (Chapter 2). Also, we excluded the cofounding factors such as familiarity of words. We reasoned that if moral self is the inner reference for information processing, then this effect should modulated by self-relatedness. Thus, if we explicitly compared the moral, neutral and immoral aspect of self and strangers, advantage of moral information should only occur on moral self. This hypothesis was confirmed by experiment 4. To further explore the interaction between morality and self-relatedness, we rendered the morality (experiment 5) or self-related (experiment 6) information as task-irrelevant information, and the interactions were found, suggesting that the moral self could be implicitly modulate the perceptual processing. To validate the stability of the effect , we conducted an meta-analysis of 6 studies, and found that the effect size of the modulation of moral self on perceptual processing is about Cohen's d = 0.5, a moderate effect size.

To further explore the mechanism underlying the facilitation effect of moral self, the current study employed cognitive modeling methods. We re-analyzed the data from the experiment 4, 5 and 6, by using drift diffusion model (DDM). The results showed that the facilitation effect of the moral-self occur primarily on drift rate, which means faster information accumulating speed. These results suggest that the perceptual salience of

moral self is similar to those of physical salience. To examine the cross-task stability of the advantage of moral self, we conducted a new experiment in which participants required to finish a perceptual matching task and a perceptual categorization task. Using ex-Gaussian model and DDM, we found that the advantage of moral self is cross task stable.

In the second part of this study, we explored the neural basis of moral self by Voxel-Based Morphometry (VBM) and meta-analysis. The VBM study explored the correlation between moral self-evaluation and volume of grey matters in the brain. We didn't found any significant cluster that correlated to moral-self. We also used activation likelihood estimation (ALE) meta-analysis of fMRI studies of moral judgment and self-referential task, because these two tasks could potentially involve moral self-referential processing. Our meta-analysis found that dorsal medial frontal cortex and frontal pole are share by both moral judgment and self-referential processing, suggesting that these two brain regions may play an important role in moral self.

In sum, the current study systematically investigated the cognitive mechanism and neural basis of moral self. From the perspective of computational modelling, we explore the mechanism of moral self form all three levels: function, algorithm and hardware. Our results showed that moral self is perceptual salient, and this saliency is cross-task stable. The cognitive mechanism behind this advantage lies in the higher information accumulation rate, and the dorsal medial frontal cortex and frontal pole maybe the neural basis of moral self. The study of moral self connected social psychology, cognitive psychology, social neuroscience, and computational modeling, providing solid foundation for future studies.

Key words: Moral self; Computational Model; Neuroimaging; Perceptual Decision Making

目　录

第一章　引言

“有两样东西，我们愈经常愈持久地加以思索，它们就愈使心灵充满始终新鲜不断增长的景仰和敬畏：

在我之上的星空和居我心中的道德法则。”

—— 康德

道德是人类的永恒话题之一，对道德的研究在哲学、社会学、法学、政治学、经济学等领域占有重要的地位。对道德的科学研究具有相当重要的意义与地位。2005 年，在 Science 杂志定义的 125 个重要的科学问题之中，有两个与道德相关的问题：“合作行为是如何演化的？”、“道德是否是大脑固有的特征？”(Kennedy & Norman, 2005)。

在科学心理学中，早期对道德的研究主要集中于发展心理学中。从皮亚杰对儿童道德判断的研究，到科尔伯格提出的道德发展阶段的理论(Kohlberg, 1969)，构成了道德心理学的主要内容。这些基于理性主义道德价值观的研究，将公平与正义（Justice）作为道德心理学研究的核心。随着女性主义以及跨文化研究的兴起，道德心理学中也开始将关爱（care）(Gilligan, 1982)和其他道德价值纳入进来(Haidt, 2007; Shweder, Much, Mahapatra, & Park, 1997)。

自 21 世纪以来，对道德的研究成为了心理学中一个热点(Greene, 2015)。尤其在社会心理学领域，对道德的研究迅速增长。在《Handbook of Social Psychology》第五版中，道德首次成为一个单独章节(Haidt & Kesebir, 2010)。此外，随着无创神经成像技术的广泛使用，如正电子发射断层显像(position emission computerized tomography, PET)、功能核磁共振成像(functional magnetic resonance imaging, fMRI)、脑磁图(magnetoencephalography, MEG)和脑电图(electroencephalography, EEG)等，社会认知神经科学结合将社会心理学、认知心理学与神经成像技术结合起来，从多个层次来研究社会心理学中的问题(Adolphs, 2003, 2009; Lieberman, 2007; Stanley & Adolphs, 2013)。在此背景之下，对道德推理的神经机制研究，也成为认知神经科学的一个重要问题(Greene, 2014)。这些来自不同角度对道德相关心理学问题的研究，大大地推进了对人们道德心理的认识，同时也提出更多待回答的问题。

1.1　文献综述

从演化的角度来讲，人类与其他动物最大的区别之一可能在于同类之间的合作。人类社会形成之后，人类的主要威胁不再是其他的物种而是同类(De Waal, 2005;

Wrangham & Peterson, 1996)。在人类社会中生存与繁荣，个体必须学会与他人合作，同时避免被欺骗带来的重大损失。有研究者认为，人类社会之所以演化出道德，主要是为了服务于合作这一功能。例如，在《Handbook of Social Psychology》第五版中，Haidt and Kesebir (2010)从功能角度对道德进行定义，将道德系统定义为：

一组相互关联的价值、美德、规范、实践、认同、体制、技术和进化出来的心理机制，它们一起抑制或者调节自私，从而让社会生活成为可能。

虽然心理学中对道德的研究一直存在，但是主要是关于儿童道德发展的研究。最近二十年来，社会心理学对道德进行了大量的探索，从多个角度对道德进行了研究，这些研究将道德的研究从单纯的道德判断与推理扩展到道德知觉、道德行为等个关键问题，研究的手段也从单一的自我报告扩展到跨文化研究、实验室实验以及神经成像等多种手段。总体而言，对道德的心理学研究在道德知觉、道德推理和道德行为三个方面取得了较大的进展。

1.1.1　道德心理学的研究现状与问题

1.1.1.1　道德作为社会知觉的基本维度

社会知觉（social perception）指的在信息加工的过程中，为了对他人的特质与意图进行准确分析的最初阶段(Allison, Puce, & McCarthy, 2000)。对于社会知觉的研究已经发现，人们对他人的社会知觉主要包括热情（warmth）和能力（competence）两个维度(Fiske, Cuddy, & Glick, 2007; 佐斌, 代涛涛, 温芳芳, & 滕婷婷, 2014)。其中热情与对他人意图的知觉有关，包括友善、乐于助人、真诚、值得信赖和道德等方面的特质；而能力则反映了对他人能力相关的特点的知觉，包括能力、技巧、创造力和效能等(Fiske et al., 2007)。先前也有研究使用道德（morality）来指热情相关的特质，但是从其定义上讲，主要是与热情相关的特质(Wojciszke, Bazinska, & Jaworski, 1998)。

由于热情与能力这两个基本维度在社会生活中的重要性，因此人们能够非常快速地对他人在这两个维度上形成第一印象。例如，Willis and Todorov (2006)发现，给被试短暂地（100 毫秒）呈现一张面孔，被试即可以对这个面孔的能力、热情等方面的进行评估，更重要的是，不同的个体对面孔的评估具有高度的一致性。Rule, Ambady, and Adams Jr (2009)则发现，即便给将面孔以不同的角度来呈现，被试也同样会得到非常准确的评估。除了根据外貌(Ambady & Rosenthal, 1992; Todorov, 2008; Todorov, Mandisodza, Goren, & Hall, 2005; Todorov, Mende-Siedlecki, & Dotsch,

2013; Winter & Uleman, 1984)，个体还可以通过他人的行为表现(McCarthy & Skowronski, 2011)和声誉（reputation）(Baumeister, Zhang, & Vohs, 2004; Dunbar, 2004)来推断他人在热情和能力上的特点。例如，Everett, Pizarro, and Crockett (2016)让被试根据他人面对道德两难情境时的决策对他人的特质进行评估，发现被试认为作出义务论判断的人更加可依赖，后续实验也重复了这一结果(Bostyn & Roets, 2017)。

虽然有研究者认为能力与热情包括了大部分社会知觉的内容，并将作为热情这个维度的一部分，但最近的研究指出，道德信息是独立于热情的另一个基本维度。例如，Goodwin, Piazza, and Rozin (2014)让人们将词从 11 个维度进行评估，发现道德品质与热情有较大的交叉，却是可以相互分离的。更重要的是，在许多情境之下，道德品质在社会知觉中具有更加重要的作用。实际上，道德不仅对于人际知觉来说至关重要，对于群体之间的知觉也起到了非常重要的作用。Brambilla, Sacchi, Rusconi, Cherubini, and Yzerbyt (2012)让被试对陌生的外群体进行评估，操纵了外群体在道德、热情和能力三个方面的特点，结果发现，道德方面的信息对于印象的影响最大。这可能是由于道德水平能够最好地反映对外群体的意图，从而影响到人们对外群体的态度与行为(Brambilla, Hewstone, & Colucci, 2013)。

总之，对他人道德信息的知觉在个体的生活占有重要地位，可能决定个体在社会生活中的成败。实际上，有研究表明，人们对他人的道德水平的知觉确实是准确的(J. J. Jordan, Hoffman, Nowak, & Rand, 2016)。正因如此，道德与热情、能力一起，形成了社会知觉的基本维度。

1.1.1.2　道德决策与道德反应

道德决策（moral decision-making）指个体面临着复杂的道德两难问题时应该如何决策；道德反应（moral reaction）则指个体面对违反道德行为时的反应(Monin, Pizarro, & Beer, 2007)。对道德判断的研究通常使用道德两难情境，例如电车两难问题：

一辆失控的电车正冲向轨道前方的 5 个工人，如果不控制电车的话，这 5 个工人就会被电车撞死。能救这 5 个人的唯一方法是扳一个开关让电车转到另一个轨道上去，但是，如果这样做，电车会撞死站在另一轨道上的 1 个人。请问，你会将电车转到另一个轨道上去，牺牲 1 个人而挽救 5 个人吗？

而对道德反应的研究，则通常使用一些明显违反了道德准则的行为，例如：

“一对兄妹，在双方同意的条件下，进行了避孕措施后，发生了性行为。他们对此保密，并且再也没有发生过性行为。请问他们的性行为是否适当？”

在心理学的研究中，对道德决策与道德反应均使用道德判断（moral judgment）

这个术语。从这两类研究的材料中可以看出，这两类心理过程的差异较大。在道德决策的研究，面临道德两难困境或者类似的情况，个体必须认真思考才得到做出决策。在这种情况下，道德决策的过程通常是慢速的、有意识的以及受到控制的。Greene, Sommerville, Nystrom, Darley, and Cohen (2001)使用 fMRI 研究人们在进行道德两难判断时的大脑活动，发现人们进行义务论判断（即认为无论何种情境都不应该以人的生命作为某种达到其他目的的工具）时内侧前额叶和后扣带回会激活更强，而进行功利主义判断（即认为可以为了更多人的生命或者幸福而牺牲少数人的生命）时背侧前额叶激活更强。由于内侧前额叶参与情绪任务，而背外侧前额叶更多地参与认知推理，因此 Greene 等推断，在义务论判断中，情绪的作用更大，而在功能论的判断中，认知推理的作用更大。

在研究违反道德准则行为的道德反应时，人们可能不需要经过太多的思考，情绪与直觉道德会指向某种反应，而认知推理的作用可能在于将该反应的结论进行合理化（justification）(Haidt, 2001)。对道德反应的神经成像研究发现，道德反应本身也并不简单，其中涉及到复杂的大脑网络的参与，包括长时记忆网络、社会知觉网络以及动机网络等(Moll, Zahn, de Oliveira-Souza, Krueger, & Grafman, 2005)。

在道德心理学中，关于认知与情绪的作用仍然存在激烈的争论。这种争议可以看作是对哲学中情与理在道德判断中作用争论的继续(喻丰, 彭凯平, 韩婷婷, 柴方圆, & 柏阳, 2011)。但另一方面，这也可能反映了由于个体面对不同情境的不同反应：面临道德两难这种比较复杂的情境时，个体需要通过认真的推理来进行权衡，而面对简单的道德违反行为时，个体则仅仅根据自己长期习得的道德准则进行反应(Monin et al., 2007)。

除了对道德两难与道德违反行为进行研究之外，还有不少研究者关注作为与不作为、动机与结果和是否有身体接触对道德判断影响(Cushman, Young, & Hauser, 2006)、道德的文化差异(Graham, Haidt, & Nosek, 2009; Haidt, 2007)等。近年来，关于道德决策，研究者又提出它是一种"心灵知觉（mind perception）"的结果，即我们是否会认为他人具有"心灵"，是否具有主体能动性，以及是否有能够感觉伤痛的受害体验(Wegner & Gray, 2016)。

虽然目前道德心理学中已经有了大量关于道德判断与道德反应的研究，但是这些研究一般使用采用假想的情境，与真实生活中的道德事件有较大的差距。因此有研究者认为，应该增强道德研究的生态学效应，更多地关注日常生活中出现的道德行为，关注道德行为的主体对自己行为的认知与判断(Hofmann, Wisneski, Brandt, & Skitka, 2014; Monin et al., 2007)。

1.1.1.3　道德行为

对道德进行研究的目的在于理解人类与道德相关的行为（以下称为道德行为，其包括道德上积极的行为，也包括道德上消极的行为）。虽然人们在日常生活中每天都能从新闻中了解到道德上消极或者积极的行为，但对真实的道德行为的研究却相对较少。

Knutson et al. (2010)给被试呈现与道德相关的关键词，让被试回忆自己做过的与这些词有关的行为。通过对这些情境的整理，作者得到了 300 多个真实的道德情境。对这 300 多个情境的因子分析表明，这些行为可以分为三个维度：准则违反（norm violation）、社会情绪（social affect）和意图（intention）。其中，最常见的不道德行为包括说谎/欺骗/不忠诚、伤害他人及拿走他人东西等。由于 Knutson et al. (2010)的研究采用数据驱动的方式，而且意在收集道德情境以作为今后道德心理学研究的材料，因此未能对全部的不道德行为进行准确地概括。

Hofmann et al. (2014)使用智能手机对美国和加拿大的 1252 名个体进行了三天的调查，每天在随机的 5 个时间点询问被试的道德行为。作者们采用道德基础理论(Graham et al., 2012; Graham et al., 2011)对这些行为进行分类，发现最多的行为属于关爱/伤害领域，其次是公正领域，再次是欺骗/诚实。

这些研究与人们的日常直觉相符合，即在真实生活情境中，出现较多的是欺骗行为。为了研究这些行为，研究者采用实验情境对这些行进行了探索。例如，Gino, Ayal, and Ariely (2009)对实验情境下的欺骗行为进行了探索。在实验中，给被试 10 美元以及 20 个包括两位小数点数据的矩阵和答题纸，被试要求在 5 分钟之内尽可能多地在每个矩阵内找到相加为 10 的数字。实验前告诉被试每在一个矩阵中找到这两个数，即可以获得 0.5 美元。5 分钟中，要求被试上交答题纸以及且根据答题情况来决定返还钱。研究者设置了两种实验条件：在基线条件下，被试需要将答题纸上交给主试；在可以欺骗的条件下，被试只需要将答题纸放到碎纸机中粉碎并将未得到的钱放到信封中后放入到一个纸箱即可。他们的结果表明，在可以欺骗的条件下，被试返还的钱数显著少于基线组。Greene and Paxton (2009)的研究也表明，个体在有机会进行欺骗的时候会进行欺骗。这些研究表明，现实生活中不道德的行为也经常发现。

综合上述关于道德知觉、道德判断与道德行为的研究，我们有如下几点发现：第一、道德信息是社会知觉中基本的维度之一，在人际知觉中起着重要的作用；第二、人们能够迅速地对他人违反道德的行为作为判断，判断他人行为是否道德；第三、在实际的生活中，人们会做出违反道德准则的行为。这些研究结果带来了一个

非常明显的矛盾：为什么人们在判断他人的道德行为时有着较高的道德标准，而自己仍然可能做出这种不道德行为？这种道德判断与道德行为之间矛盾的一类现象是道德虚伪（moral hypocrisy）(Batson, Kobrynowicz, Dinnerstein, Kampf, & Wilson, 1997; Batson, Thompson, Seuferling, Whitney, & Strongman, 1999)。正是由于道德判断与道德行为之间不匹配，引出了一个重要的问题：道德判断与道德行为中的冲突中，个体如何维持一个统一的自我概念？这就涉及到道德心理学中一个探索较小的问题：道德自我。

1.1.2　道德自我的概念辨析

1.1.2.1　道德自我的概念

道德心理学研究长期关注于道德判断，但道德判断与道德行为之间不一致，导致道德判断的研究停留在理论层面。一些研究者探索了这种不一致的原因，他们发现，道德自我（moral self）可能起着非常重要的作用。Blasi 对道德自我的作用进行了开创性的说明，指出自我认同（identity）在道德判断与道德行为（moral action）关系起到调节作用(Blasi, 1984, 2004)。此后，研究从不同的角度对道德自我进行探索，这些研究表明，道德自我既具有稳定性，也具有情境性(Monin & Jordan, 2009)。

道德自我的稳定性表现为道德自我具有与人格类似的特点。研究者从道德认同（moral identity，也翻译为道德同一性）的角度来定义道德自我。例如，Blasi 将道德认同定义为言行具有道德是个体自我本质（essential self）组成部分的程度(Blasi, 1983)。与这一定义相似，Aquino and Reed II (2002)将道德认同定义为个体对那些道德原型所具有特质的认同程度。在研究中，他们定义出道德原型的核心特质后，让被试评估拥有这些特质对自己的重要性（内化维度，internalization）以及在与他人的交流中的重要性（外在的符号化，symbolization）。这种道德认同与自我价值(Crocker & Wolfe, 2001)中的美德对自我价值的影响有着类似的影响。此外，也有研究者从道德图式（moral schema）或者的角度来定义道德自我，即个体在多大程度上从道德的角度来对外界信息进行编码。由于道德图式是具有长期的特点，因此有研究将其称为道德的长期性（moral chronicity）(Narvaez, Lapsley, Hagele, & Lasky, 2006)。

也有研究者从情境性和可变性的角度来定义道德自我。Monin and Jordan (2009)提出道德自评（moral self-regard）的概念，认为道德自我是个体即时对自己道德程度的评估。虽然 Monin and Jordan (2009)承认道德自评的稳定性，但他们更强调道

德自评随情境因素的变化而具有动态的特点。此外，道德自评作为自我概念的一部分，具有强烈的动机性。这种动机性主要表现为人们需要维持积极的自我形象（self-image），尤其是积极的道德自我形象。这种积极的道德自我形象对个体来说非常关键，甚至可能用来弥补个体在其他领域的失败(A. H. Jordan & Monin, 2008)。与道德自评相似的概念是道德自我形象（moral self image），也是指个体易变的道德自我概念（moral self concept），即他们自我概念中与道德原型特质相关的部分(J. Jordan, Leliveld, & Tenbrunsel, 2015)。

事实上，人们不仅认为道德自我具有重要的意义，而且对道德在他人身上的重要性也具有同样的认识。Strohminger and Nichols (2014)比较了人们心理与认知功能的变化对身份同一性（Identity）的影响，发现道德品质（moral traits）的变化会让被试觉得人们不是是原来的那个人，而其他方面的变化如感觉知觉或者记忆能力的变化则不会影响身份同一性。通过让被试判断某个人在经过一些变化之后是否仍然是原来的“他”。Molouki and Bartels (2017)进一步发现，个体在道德品质向消极方面的变化会导致人们知觉到的身份同性的变化，但是向积极方面的变化则不会。这表明，个们倾向于认为积极的道德品质是每个个体所固有特征。确实，Newman, Bloom, and Knobe (2014)发现，人们倾向于认为每个人的真我（true self）都是道德的。

基于以上道德自我的相关研究，本文认为，道德自我（moral self）是人们在道德方面的自我概念（self-concept），其是自我概念的核心组成部分。道德自我既包括个体长期以来形成的对自己道德水平的评估，也包括个体在具体情境下对自己道德水平的即时评估。道德自我具有强烈的动机性，维持积极的道德自我概念对于个体来说具有重要的意义，因此对道德自我的威胁会带来个体特定的反应。

1.1.2.2 道德自我与其他自我概念的关系

道德自我作为自我概念中的一部分，与先前研究中关于自我概念的定义有着密切的关系。首先，道德自我属于心理自我（psychological self）的内容，与身体的自我相对。虽然心理学家，如 William James，对心理自我进行更加细致的区分，认为自我包括身体自我、社会自我与精神自我(James, 1890)，但是在心理学的研究中，使用心理自我与身体自我的两分法也比较常见(Hu et al., 2016)。而道德自我既的内容既包含了社会自我的成分，也包括了精神自我的成分。因此，可以将个体的道德自我作为心理自我的一部分。

其次，道德自我与自我不一致理论（self-discrepancy theory）中的实际自我（actual self）、应然自我（ought self）和理想自我（ideal self）也不同。自我不一

致理论中实际自我指的是个体为他人所知觉到的我；理想自我指个体想要成为的我，而应然自我指的是个体应该成为的我(Higgins, 1987)。道德自我与自我不一致理论的区别在于，道德自我专指个体对自己在道德水平的评估，而自我不一致理论中的是一般性评估。此外，自我不一致理论中，将个体自己对自己的评估与知觉到的他人对自己的评估区分开来，但是道德自我并不区分这两种评估的来源。

再次，道德自我与真我（true self）也有区别。真我也叫内在自我（intrinsic self），指的是人们认为自己真正是谁(Bargh, McKenna, & Fitzsimons, 2002; Landau et al., 2011)。由于人们可能因为各种外界的原因而无法表现出自己真实的一面，因此真我的研究对于了解人们的自我概念也具有重大的意义。虽然最近的调查发现，人们倾向于认为真我的是道德的(Newman, Bloom, et al., 2014; Newman, De Freitas, & Knobe, 2014)，但是与自我不一致理论中相似的是，真我也包含了道德之外的领域，例如能力。而道德自我概念中包括了人们认为是真我道德成分的部分，也包括了自我道德中不属于真我的部分。

总之，道德自我是自我概念中与道德密不可分的部分。在具体研究之中，研究者一般会对其进行操作化定义。例如在 Aquino and Reed II (2002)对道德自我认同的定义中，作者们首先确定了 9 个被认为是道德的人所具有的核心特质，然后再让被试评估自我这 9 个核心特质对自己的重要性，与此类似，J. Jordan et al. (2015)对道德自我形象的定义也采用核心特质词的方法。在对道德自我情境性进行研究中，研究者一般采用道德行为测量，例如亲社会行为(Mullen & Monin, 2016)。而在本研究中，我们认为道德自我是自我概念的一部分，具有稳定性和抽象的特点，因此这里的道德自我指个体从道德方面对自己的整体评估，它既包括积极的部分（如道德高尚的一面），也包括消极的一面（如道德低下的一面），还包括中性的一面。

1.1.3　道德自我的优势及其对行为的影响

1.1.3.1 道德自我在自我概念中的优势地位

与道德相似的是，对于自我与身份（self and identity）的研究是社会心理学中最核心的问题之一(Swann & Bosson, 2010)。在对自我的研究中，研究者们发现，人们对自我的认识往往并非客观和真实，而是带有许多的偏见，这些偏见表现在认知的各个层面。其中最具有代表性的效应，即自我增强效应（self-enhancement effect），即人们会认为，与自己所在群体的平均水平相比，自己会更加聪明、有吸引力、健康，自己的生活将更幸福、寿命更长等。这个效应也被称为优于平均水平效应

（Better than average effect）(Alicke & Govorun, 2005)。自我增强的效应与个体的幸福感有密切的关系(Taylor & Brown, 1988)，人们使用这种积极的视角来看待自己，可能正是为了维持一种积极、下面的自我形象，从而能够应对逆境。

自我增强效应在道德水平也存在，人们普遍认为自己的诚实和利他的水平高于平均水平(Pronin, 2008)。与自我概念其他方面不同的是，道德自我概念的增强效应似乎比其他领域更具有优势。例如，Epley and Dunning (2000)发现，人们对于道德的行为，更倾向于说是自己去做的，但是对于不道德的行为，则会更少。这一结果最近也得到了部分的重复(Klein & Epley, 2016)。Brown (2012)发现，虽然被试普遍存在着优于平均水平的效应，但是在道德领域相比于能力而言，这种效应量更大(Brown, 2012)。在道德与学业水平上，也存在着同样的效应(Möller & Savyon, 2003)。此外，道德方面的自我增强效应比智慧(Zell & Alicke, 2011)、野心(Alicke, Vredenburg, Hiatt, & Govorun, 2001)、和智力(Van Lange & Sedikides, 1998)方面均更加强。

最近，Tappin and McKay (2016)对自我在道德领域的这种优势效应进行了进一步探索。他们让被试对一批特质词进行评估，判断每个词是否适合形容自己、是否适合形容其他人的平均水平以及是否是社会所赞许的特质。这些特质词分成了道德、能力和热情三个维度。他们的结果表明，人们更加倾向于认为道德方面积极的词汇适合于描述自己，而且这种得分要高于其他方面的自我增强。

这些研究结果与前述道德自我对于个体身份同一性的结果一起，均表明道德自我在自我概念中占据有核心的地位。道德自我增强上更强的效应，表明道德自我对于个体来说具有更重要的地位。道德自我在自我概念中的核心位置，与道德在社会知觉中的基本维度的地位一起，表明道德在社会中的重要作用，不仅体现在人与人之间的关系，也表现在个体内部。

1.1.3.2 道德自我对行为的影响

由于道德自我在自我概念中的核心地位，人们也具有强烈的动机要维持积极的道德自我形象。因此，道德自我的变化，能够对随后的行为产生影响。

已有的研究表明，道德自我对行为的影响主要表现为如下三个效应：道德许可（moral license）、道德憎恨（moral resentment）和道德补偿效应（moral compensation）(Monin & Jordan, 2009)。其中道德许可效应指的是当个体的积极的道德自我概念得到保障之后，他们可能会做出一些不道德的行为(Effron & Conway, 2015; Mullen & Monin, 2016)；道德憎恨指的是当个体的积极道德自我受到威胁时，他们可能会憎恨威胁的来源，即便这些威胁来源是道德高尚的榜样(Monin, Sawyer, & Marquez,

2008)；道德补偿指的是个体在道德之外的领域受到威胁之后，会加强自己的道德自我以达到补偿的效果(A. H. Jordan & Monin, 2008)。此外，积极的道德自我概念受到威胁，人们会忽视道德标准或者变化自己的道德标准来维持积极的道德自我(Mazar, Amir, & Ariely, 2008)。

1.1.4　道德自我对认知加工的影响

虽然社会心理学中对道德自我的特点以及道德自我对于道德行为的影响上进行了大量的研究，但是道德自我如何影响道德行为的机制却仍然不清楚。作为一种高级的社会认知，其对行为的影响可能始于低级的认知过程，因此了解道德自我对认知过程的影响对理解道德自我具有重要的意义。虽然目前没有研究者对道德自我是否能够对低级的认知过程产生影响进行探讨，但是基于先前的研究，我们推断道德自我能够对认知加工产生影响。具体而言，从心理机制上讲，道德自我可能具有社会突显性；从神经机上讲，道德自我可能与默认网络有关，通过突显网络来调节认知加工。

1.1.4.1　心理机制上的可能性：社会突显性

从信息加工的角度来讲，人类的认知资源有限，必须将认知资源优势分配到有价值的信息之中。因此，道德自我信息作为一种重要的信息，在认知加工上应该具有与其他信息不一样的优势。

认知心理学的研究表明，在知觉加工水平上即可得到优先加工的刺激是那些具有知觉突显性（salience）的刺激。知觉的突显性分成两类：物理的突显性（physical salience）和社会的突显性（social salience）。前者指的是刺激物理特征，如明度、对比度等，与低对比度的刺激相比，高对比度的刺激能够迅速被注意到。后者则是指由于刺激对人意义带来的突显性，例如自我相关性或者情绪的作用。例如对人类祖先来说具有威胁性的蛇和蜘蛛(Öhman, Flykt, & Esteves, 2001)，而且甚至当这些刺激不是我们所要注意的目标时，他们也会吸引我们的注意(李会杰 & 陈楚侨, 2007)。已有的研究已经表明，对于个体生存来说非常重要的信息，会因此变得具有突显性，例如情绪刺激(de Gelder, van Honk, & Tamietto, 2011)和奖赏(B. A. Anderson, Laurent, & Yantis, 2011)相关的信息确实能够得到优先的加工。由于道德自我的重要性，人类的感觉知系统可能会对这类信息进行优先的加工(Cosmides & Tooby, 2000)。

根据已有的研究，道德自我相关的信息也可能具有认知加工上的优势。首先，

对自我参照的研究表明，自我参照对认知的影响是多个层次的，自我参照的过程能够像“胶水”一样，在知觉、注意、记忆和推理等多个层面上，将信息进行绑定从而进行快速的加工(Sui & Humphreys, 2015a)。早在 1977 年，研究者就发现，当要求被试在记忆任务中，想像呈现的词语是否适合形容自己时，记忆的效应要优于其他编码方式(Rogers, Kuiper, & Kirker, 1977)，后续的研究发现，这个效应非常稳定(Zhu, Zhang, Fan, & Han, 2007)。在意识方面的研究也表明，自我相关的信息可以优先进入意识(Macrae, Visokomogilski, Golubickis, Cunningham, & Sahraie, 2017)，失去意识的患者在听到自己的名字时，脑电上出现了响应(Di et al., 2007)。最近在知觉方面的研究表明，不管是自我面孔(Ma & Han, 2010; Sui & Humphreys, 2013)、自我名字(Nakane, Miyakoshi, Nakai, & Naganawa, 2016)还是仅与自我相联系的几何图形(Sui, He, & Humphreys, 2012)，均在知觉加工中具有优势。例如，Sui 等让被试在实验室中学习几何图形与不同人物标签之间的关系（例如，正方形代表你自己、三角形代表你最好的朋友、圆形代表一个陌生人），随后让被试快速地判断图形与文字是否匹配，结果发现，与自我相联的图形条件下，被试的匹配判断显著快于朋友或者陌生人联系的图形(Sui et al., 2012)，表明自我具有高度的社会突显性。更重要的是，自我的社会突显性会与刺激的物理突显性发生交互作用：与自我相关的图形会有更强的突显性，从而使用原先不具有突显性的局部刺激得到优先加工(刘明慧, 张明, & 隋洁, 2014)；或者使得原本加工更快的小图形不再有优势(Sui & Humphreys, 2014)。

其次，已经有研究表明，道德相关的信息也在认知加工中具有优势。Gantman and Van Bavel (2014)使用道德、非道德词汇以及假词作为刺激，短暂呈现（例如 33ms）刺激后立刻呈现掩蔽刺激，让被试判断目标刺激是词还是非词。结果表明，对于道德相关的词汇，被试的判断准确率更高。进一步的研究表明，道德词汇的这种优先加工效应，可能是由于被试的道德动机（维持正义）所导致，当采用阅读道德故事的方法让被试的正义动机得到满足之后（比如让被试阅读一段坏人得到惩罚的故事），道德词汇的优先效应消失(Gantman & Bavel, 2016)。此外，Shore and Heerey (2013)将中性面孔与不同的合作历史进行匹配之后，让被试进行注意瞬脱（attention blink）任务，结果发现，如果该面孔有着不合作/欺骗的历史，则被试的注意瞬脱效应会减小。E. Anderson, Siegel, Bliss-Moreau, and Barrett (2011)通过双眼竞争实验，发现对于同样不带明显情绪的中性面孔，那些被赋于了不道德行为的面孔在我们的视知觉中更占主导。

再次，在记忆方面的研究表明，人们对自己道德相关的行为的记忆是具有选择性的。Kouchaki and Gino (2016)让被试假想自己做出道德的和不道德的行为，几天

后测量其是否仍然对这些行为有较好的效果，结果发现，被试对道德行为的记忆效应要明显优于不道德行为。其他的研究者也发现了类似的效应(Sedikides, Green, Saunders, Skowronski, & Zengel, 2016)。这可能反应了记忆的对道德自我的保护作用(Sedikides & Green, 2009)。

以上三方面的研究表明，自我相关信息和道德信息均在知觉加工中具有优势，而道德自我也在记忆方面表现出加工优势。这些研究表明，作为自我概念中的核心部分，道德自我可能在各个认知加工的过程中均有重要的影响。

1.1.4.2　神经机制上的可能性：默认网络

近年来对自我的神经成像研究表明，与心理自我相关的信息在皮质中线结构（cortical midline structures, CMS）的脑区加工，其中最主要的是腹内侧前额叶、前扣带加、扣带回中部（mid cingulate cortex, MCC）和后扣带回(Hu et al., 2016; Northoff et al., 2006)。事实上，心理自我加工的网络与默认网络（default-mode network, DMN）具有非常大的重合度，例如前扣带回在进行自我相关的加工时激活，在被试静息状态下也激活(Qin & Northoff, 2011)。而且，前扣带回与在自我特质上的个体差异有关(Wang et al., 2016; 王非, 2015)。正是由于自我信息加工网络与默认网络的相似性，有研究者认为两者是相同的，认为 DMN 可能反映了个体自发的大脑活动，与自我加工有关。

对于道德信息加工的 fMRI 研究也表明，默认网络中的重要脑区如 PCC 也参与道德加工。例如 Greene et al. (2001)发现，当被试面临道德两难困境时，当他们做出符合义务论的道德判断（即认为任务时候不得将人的生命当作工具）时，PCC 和 mPFC 激活强烈。Harrison et al. (2008)发现，当被试面临道德两难情境时，他们的默认网络比静息状态下激活更加强烈。而 McDonald et al. (2017)也重复了这一结果。这表明，在道德推理中默认网络起到了重要的作用。由于在道德两难情境中，被试没有外显的道德标准可以参考，因此需要以自己的道德价值为参考。从这个角度来讲，道德两难情境下，可能是一种内隐的道德自我加工。

默认网络指的是在静息状态下激活而进行认知任务时出现负激活的大脑网络(Biswal, Yetkin, Haughton, & Hyde, 1995; Greicius, Krasnow, Reiss, & Menon, 2003)。对于 1000 多个被试在静息状态下的功能核磁共振成像研究表明，DMN 是大脑皮层内在联结的网络之一(Yeo et al., 2011)。典型的 DMN 包括后扣带回（posterior cingulate cortex, PCC）、前扣带回（anterior cingulate cortex, ACC）和内侧前额叶（medial prefrontal cortex, mPFC）。

默认网络是目前认知神经科学关于神经网络研究中最重要的一个。这些网络

是大脑在完全静息状态下和任务状态下，均有一些同时激活的大脑区域，它们可能是人类认知与行为的基础(Park & Friston, 2013)。对这些神经网络的研究表明，大脑内存在一些功能上不同的网络，这些网络在静息状态下和在进行任务加工时具有相似性(Cole, Bassett, Power, Braver, & Petersen, 2014; Yeo et al., 2011)，并且不同的神经网络对应着不同的认知功能。

在已知的神经网络中，默认网络与自我信息的加工以及道德推理均有较大的重合，这可能构成了道德自我的神经基础。而默认网络与大脑内与认知加工相关的神经网络如突显网络（salience network, SN）和中央执行网络（central executive network，CEN）有着密切的关系，可能是道德自我能够对其他认知加工产生影响的神经基础。

突显网络指用来探测外界刺激和体内信息突显性并将注意朝向该刺激的神经网络(Seeley et al., 2007)。突显网络主要包括背侧前扣带回（dorsal anterior cingulate cortex, dACC）和眶额皮层-岛叶结构（orbital fronto-insular cortices）。突显网络与皮层下结构和边缘网络有着广泛的神经联结，并且其与中央执行网络是相互独立的(Seeley et al., 2007)。

中央执行网络指用负责高级认知功能如计划、决策等功能以及控制注意与工作记忆等认知功能的大脑网络(Menon, 2011)。中央执行网络主要包括背外侧前额叶和顶叶的新皮层(Seeley et al., 2007)。

虽然目前没有关于道德自我的神经成像研究，但是心理自我的研究与道德两情境的研究均表明，默认网络参与到道德推理与自我信息的加工。由于道德自我是心理自我的核心部分，而道德两难情境可能涉及到内隐的道德自我参照加工，因此道德自我可能会涉及到默认网络。

如果道德自我的神经基础在于默认网络，那么其对认知功能的调节作用也成为可能。已有的研究表明，默认网络与中央执行网络负责不同的功能，但是突显网络可能起到调节作用。Goulden et al. (2014)发现，突显网络在 DMN 与执行网络之间的切换起着重要作用。在道德推理过程中，Chiong et al. (2013)发现，突显网络能够调节到默认网络的活动。在自我信息的加工中，Sui, Rotshtein, and Humphreys (2013)发现，vmPFC 对与额顶中央执行网络的共同激活可能是自我信息加工产生优势的原因，而 vmPFC 正是默认网络的关键脑区。

以上的分析表明，虽然目前尚未有研究直接探讨道德自我会影响认知加工的神经基础是什么，但从已有的发现中，我们可以推断：默认网络，尤其是前扣带回/内侧前额叶，可能是道德自我的神经基础，而默认网络与突显网络和中央执行网络的相互关系，可能是道德自我在认知加工中具有优势的神经基础。

1.2　问题的提出

1.2.1　研究现状总结和本课题的研究问题

如前所述，道德如何指导个体行为是该领域研究的核心问题之一。对道德自我的研究主要集中于道德自我的特性（例如，道德自我的情境性或者稳定性(J. Jordan et al., 2015; Monin & Jordan, 2009)）、道德行为的关系(Aquino & Reed II, 2002; Mullen & Monin, 2016)以及道德自我在自我概念中的优势(Tappin & McKay, 2016)，却没有研究系统地讨论道德自我的心理与神经机制。只到最近，研究者才开始关注道德自我对认知过程的影响，发现道德自我能够调节对道德事件的记忆(Kouchaki & Gino, 2016)。但是，没有研究系统地讨论道德自我是否在认知加工中的各个阶段上均有影响。

正如 1.1.4 小节中所分析，道德自我对认知过程的影响在行为上和神经上均是具有可能性的。但目前这种可能性是否确实存在，以及其内在的认知机制与神经机制到底是什么，仍然缺乏系统研究。

因此本课题采用行为实验、元分析和神经成像技术来探索道德信息影响知觉的认知机制和神经基础。一方面，使用系统和严格的实验室实验来探索和验证道德自我对知觉的调节作用，并且使用已有的计算模型去探索这种调节作用背后的认知机制；另一方面，使用大样本的神经成像数据，探索道德自我的神经基础。

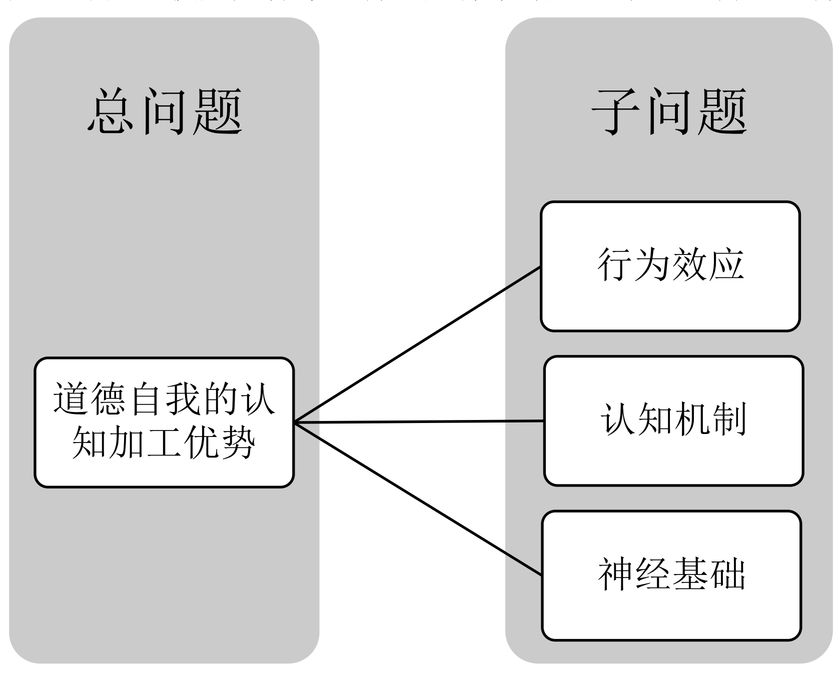


图 1.1　本研究的总问题和两个子问题

1.2.2　研究框架

道德自我是个体道德与自我概念的核心部分，探索道德自我在知觉加工的调节作用，本质上是对道德自我在认知过程中的优势作用。本文研究在对其探索的过程中，采用了如下思路：首先，采用行为实验和元分析的方法，来验证道德信息与道德自我对知觉加工的影响；其次结合计算模型与行为实验的方法，探索道德自我对知觉加工影响的认知机制；最后采用神经成像的方法，探索道德自我的神经基础（见图 1.2）。

在行为实验中，为了能够将道德自我这个抽象的概念与知觉加工的任务结合起来，分离出道德自我与非道德自我在知觉上的差异，本课题采用如下三个标准来选择行为实验的范式：（1）实验刺激简单明确，能够进行简单的知觉任务，且易于调控；（2）实验任务简单易懂，个体能够在实验室中掌握；（3）实验刺激本身尽量少地受到混淆变量的影响。

基于以上三点考虑，本研究中选择了社会联结范式（social associative paradigm）(Sui et al., 2012)。在这个任务中，被试需要学习几个简单几何图形与不同人物之间的对应关系，建立起联结。当这种联结关系建立起来之后，可以以几何图形为目标进行其他认知实验，例如，知觉匹配任务、视觉搜索任务等。由于社会联结范式采用简单的几何图形作为刺激，因为刺激本身简单明确，且避免了先前研究中使用面孔带来的自发特质推断（spontaneous trait inference）(Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015)带来的混淆，其联结学习的过程能够迅速在实验室中完成，符合实验室实验的特点，因此本研究在探索道德自我在知觉过程上的加工优势时，以社会联结范式作为基本方法，考察道德自我在知觉过程中是否存在优势以及这种优势的认知机制。

以社会联结范式为基础，我们首先考察了道德信息是否在知觉加工中具有优势作用。由于先前的研究已经表明，自我参照加工在知觉中会表现出非常稳定的优势(Sui et al., 2012; Sui & Humphreys, 2015b)。相比而言，道德信息是否在知觉加工中具有优势由受到争议。虽然 Gantman and Van Bavel (2014)采用词汇判断任务时，发现道德信息具有优势，但是这一现象可能是由于语义激活所导致(Firestone & Scholl, 2014)。因此，在对道德自我的知觉加工优势进行探索之前，我们验证了道德信息在知觉中加工的优势是否存在。

具体而言，我们采用操纵了道德品质信息的效价，使用三个分别代表道德上积极、道德上中性和道德上消极的人物标签（好人、常人和坏人）来代表积极、中性和消极的道德品质信息（第二章）。之所以选择道德品质信息，是因为道德品质信息是我们在社会知觉中研究最基础的维度（见本章 1.1.1.1 小节），而且对其知觉

有快速和自发的特点，适合知觉研究。使用三个标签词代表三种不同的道德品质信息，能够让被试快速地学习到道德信息的效价，而避免了道德情境的个体差异带来的实验操控的困难。

为探索道德自我对知觉加工的影响，我们在道德品质信息的基础之上，进一步加入自我参照这个变量。具体而言，我们一方面直接比较了道德自我（“好我”）与非道德自我（如普通水平的我：“常我”，和不道德的自我：“坏我”）在知觉加工中的差异，另一方面，还进行了内隐的任务，让道德信息或者自我信息变成与任务无关的信息时，考察它们对知觉加工的影响（第三章）。更重要的是，我们通过元分析的方法，检验道德自我在跨研究间的泛化效应。

为了进一步考察道德自我在知觉加工优势的认知机制，我们采用认知科学中模型分析的方法，对反应时间和正确率数据进行模型的拟合。模型分析的方法，不仅能够更加充分地利用反应时间和正确率的数据，还可以对认知过程进行分解，从而探索道德自我对不同认知过程的影响。根据我们研究中反应时间的特点，我们将采用扩散模型（drift diffusion model）和指数高斯模型（ex-Gaussian model）来对反应时间进行模型分析。

在模型分析过程中，我们同时研究了道德自我跨任务的稳定性。通过结合知觉匹配任务与知觉分类任务两个任务，本研究考察了道德自我在知觉中优势加工是否在不同的任务下均能够表现出稳定的优势，从而考查其可靠性。

最后，为研究道德自我的神经机制，本研究进行了探索。一方面，通过对现有研究的神经成像研究进行元分析，找到与道德和自我均相关的大脑区域。另一方面，以元分析的结果为基础，采用基于体素的形态学分析（ Voxel-based Morphometry，VBM）方法，考察大脑区域结构特征（如灰质体积）与个体道德自我评估的关系，进一步验证特定大脑区域与道德自我之间的关系。

1.2.3　研究方案

本研究从行为层面和神经层面对道德自我进行研究。行为层面，即包括对道德自我在知觉中加工优势（即突显性）效应是否存在进行验证，在此基础上，进行一研究了其认知机制以及跨任务的稳定性。神经层面，采用元分析方法和 VBM 的方法，对道德自我相关的脑区进行功能与结构的定位。综上所述，在本课题的研究框架如下（见图 1.2）：

（1）第二章：研究 1 ~ 3：道德信息的知觉加工优势

这部分研究属于对先前研究中存在争议部分进行探索验证。通过采用社会联结范式，让被试在实验室情境之下，学习简单几何图像的道德意义，然后进行知觉匹

配任务。这一方法有效地控制了混淆变量，能够清楚地操纵道德信息的效价，从而考查道德信息对知觉的影响。此外，使用认知心理学中常用的策略，我们通过实验控制可能混淆变量的影响，重复了实验中的效应。

（2）第三章：研究 4 ~ 7：道德自我的知觉加工优势

这部分研究开始探索本研究中的核心问题：道德自我是否具有知觉上的加工优势。通过操纵几何图形与联结的词汇的社会意义，我们系统地考查了道德自我与其他控制条件之间的差异。更进一步，我们通过元分析方法，对行为实验效应进行了综合。

（3）第四章：研究 8 ~ 9：道德自我在知觉中的锚定作用

这部分研究开始深入探讨道德自我的认知过程。采用指数高斯模型与扩散模型，我们不仅对先前行为实验进行重新的分析，而且通过结合知觉匹配与知觉分类任务，研究了道德自我的知觉加工优势跨任务的一致性，从而更深入地了解道德自我在知觉过程中的哪个阶段具有优势。

（4）第五章：研究 10 ~ 11：道德自我的神经机制

这部分研究关注于道德自我相关的大脑区域，即功能定位。一方面，通过 VBM 分析，来探索与道德自我评估相关的大脑结构；另一方面，对先前心理自我和道德判断神经成像研究的元分析，找到道德与自我共同激活的脑区。

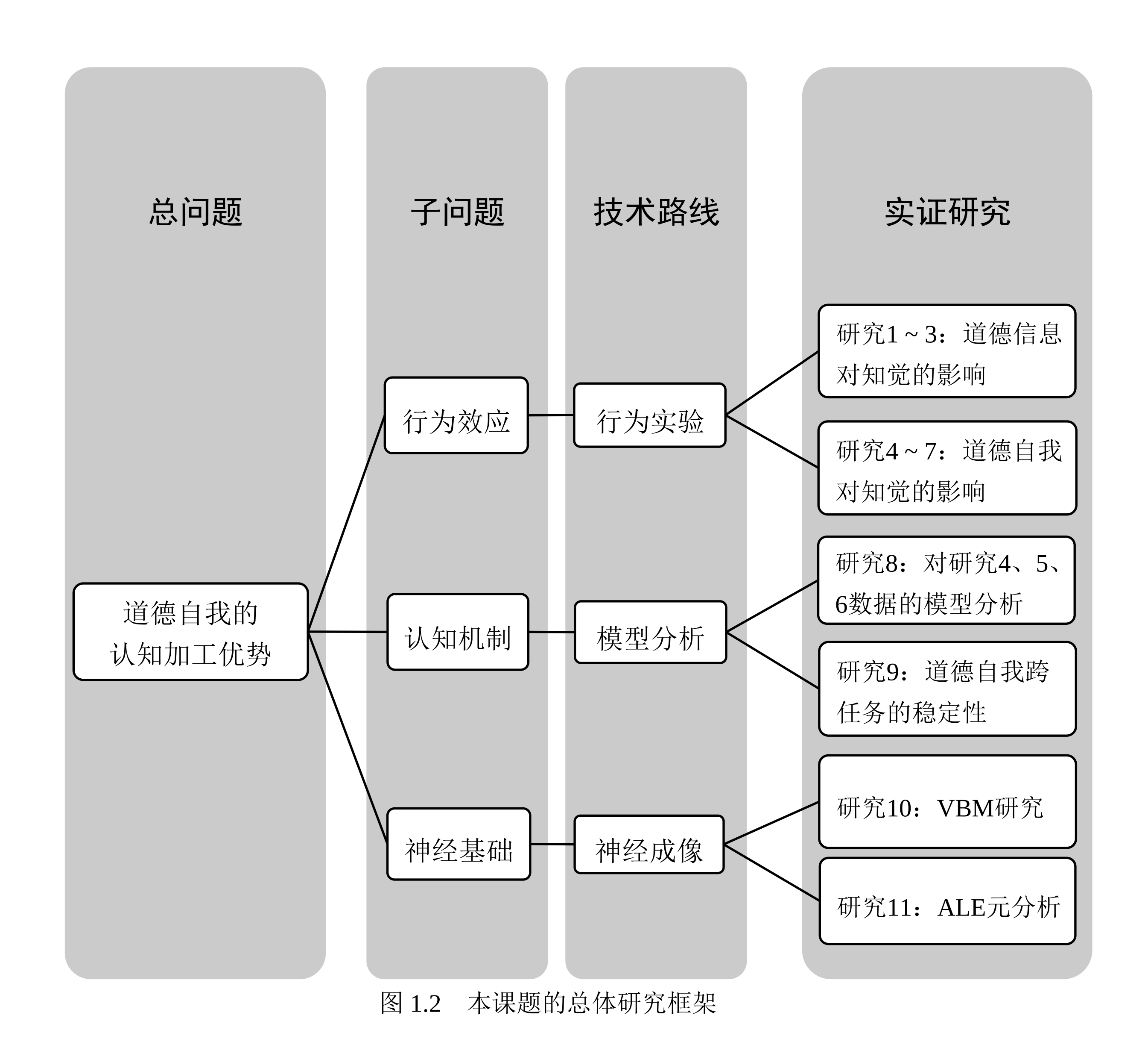


图 1.2　本课题的总体研究框架

1.2.4　研究意义与创新之处

与道德判断相比，道德自我概念更加能够预测个体的道德行为，但是道德自我对认知过程有什么影响以及其神经机制如何却并不清楚。本研究探索了道德自我对认知过程的影响及其神经机制，尝试去回答上述问题，对于今后进一步研究道德自我奠定基础。本研究的结果将有助于我们去理解人类的道德行为，以期能够进行预测和改变人类的道德自我，从而改变人们的道德行为。从计算认知神经科学的角度来讲，结合认知模型与神经科学的方法，将有助于我们进一步去理解心理过程的神经相关物。本研究将知觉决策领域的模型引入社会认知的研究，从而能够更加深入去理解社会认知过程中的心理机制问题，为今后的计算认知神经科学研究打下基础。

此外，本研究结合元分析方法与 VBM 方法，对道德自我进行了功能与结构的定位，这将为今后的功能成像研究提供借鉴。由于科学界的可重复危机，先前发表的研究是否可靠受到了质疑(Szucs & Ioannidis, 2017; 胡传鹏 et al., 2016)。元分析方法使用统计方法，从多个研究中找到稳定的结果，能够得到比单个研究更可靠的结论(Wager, Lindquist, & Kaplan, 2007; 胡传鹏, 邸新, 李佳蔚, 隋洁, & 彭凯平, 2015)，因此受到了研究者的推荐(Poldrack et al., 2017)。本研究在神经机制的研究中，使用元分析的方法找到道德自我的功能定位，其结果更稳健可靠。以元分析结构作为 ROI，对大样本的神经成像数据进行 VBM 分析，在方法也具有严谨性。这些研究将为后续的道德自我研究提供坚实的基础。

在现实意义层面，人们每天都在面临各种复杂的情境，需要作出无数的决策，而这些决策中最基本的是知觉决策。人们的注意是如何导向的，从什么视角来看待问题将直接影响到随后的认知判断。本研究表明，道德自我在知觉中具有优势，其存在使得人们更多地去对他人做出道德评判，作为“道德家”，而非“科学家”(Knobe, 2010)。这种道德自我的偏见，也可为今后的公共政策、人工智能等提供实证基础。

第二章　道德信息是否具有知觉加工优势？

2.1　本章引言

对道德的研究中，常常关注于道德判断以及道德行为（如博弈中的利他行为）。但是对道德信息的知觉加工却常常受到忽略。虽然社会心理学中有社会知觉（social perception）相关的研究，但是这些研究实质上是研究人们对他人的印象形成(Adolphs, Nummenmaa, Todorov, & Haxby, 2016)，而非认知心理学意义上的知觉。

认知心理学上的知觉，指的是更简单的对信息加工的过程。通常被定义为对刺激的整体加工。认知心理学中对知觉进行了大量的研究，但通常都是以非自然的、简单的刺激，例如光栅、几何图形。这些研究提供了大量关于人类知觉特点的知识，为理解人类认知功能提供了坚实的基础。例如，目前比较一致承认的关于知觉的基本的格式塔原则，以及关于知觉过程中的 what 通道和 where 通道。

但知觉的研究中，也发现了一些与社会信息相关的特别，例如人类对面孔以及人类对威胁性刺激的加工特点。对面孔的研究表明，人们通常能够在极短的时间内知觉到正立面孔的存在，但是对倒立面孔却没有这种优势(Jiang, Costello, & He, 2007)。而且这种知觉受到个体动机的影响(Jiang, Costello, Fang, Huang, & He, 2006)。而对情绪(Hedger, Gray, Garner, & Adams, 2016)、自我(Sui & Humphreys, 2015a)和奖励(B. A. Anderson et al., 2011)的研究也表明，人们对于具有情绪性的、与自己有关的以及具有高奖励的刺激的知觉加工更快。

这些研究表明，对于个体具有社会意义的刺激，能够在知觉加工中被优先加工。如前所述，道德信息对于个体而言，具有重大的社会意义。尤其是关于他人道德品质的信息，可能直接关系到个体在合作中是收益还是损失。因此对于道德品质信息的加工，从理论上应该能够在知觉加工中具有优势。

虽然已经有少数的研究表明，道德信息对人类知觉信息的影响可能在很早就开始了。例如 E. Anderson et al. (2011)发现，将中性面孔与不同的社会信息匹配之后，负性情绪匹配的面孔会在视觉的双眼竞争中占主导的时间更长。类似的是，Shore and Heerey (2013)发现，将中性面孔与不同效价的社会信息匹配之后，会减少注意瞬脱（attention blink）效应。然而，这两个研究中均采用了面孔作为实验材料，而面孔本身很容易引起人们对其道德特质的自发推理(Todorov et al., 2015)。

Gantman and Van Bavel (2014)通过词汇判断的任务，让被试判断短暂出现并且被掩蔽的刺激是词还是非词，结果发现对于道德词汇的识别准确率更高。但是 Firestone and Scholl (2014)却通过实验表明，由于道德词汇之间的语义联系，对道

德词汇更加准确的识别可能反映了一种语义上的启动，而非道德本身的知觉加工优势。

由于先前这些研究中局限性，道德信息是否具有知觉上的突显性仍然是一个有争议的问题。为了验证道德信息的突显性，本研究采用三个行为实验。为了控制面孔带来的复杂信息以及使用词汇带来的语义启动问题，我们采用社会联结学习范式。该范式中，需要将学习简单的几何图形所代表的社会意义，然后进行简单的知觉匹配任务。使用该任务，我们通过三个行为研究来验证道德信息的知觉加工优势。

首先，研究 1 对道德信息的知觉加工优势进行初步的探索。如前所述，我们操纵了道德品质信息的效价，使用三个分别代表道德上积极、道德上中性和道德上消极的人物标签（好人、常人和坏人）来代表积极、中性和消极的道德品质信息。然后让被试将几何图形与不同道德品质的人进行联结，我们检验了积极、消极和中性的道德品质对知觉匹配任务的影响。

其次，为了排除词汇本身的影响，研究 2 中选取具有同样意义但是更加不常用的词作为刺激，来对道德突显性进行重复。

最后，研究 3 采用先后呈现的方式，让被试先看到词汇，然后仅对图形做反应。通过后掩蔽的方式消除了词汇的物理性的影响之后，考查图形本身对知觉加工的影响。

2.2　研究 1：道德信息是否具有知觉加工优势？

2.2.1　引言

研究 1 的主要目的是为了验证道德信息是否具有知觉加工优势。

2.2.2　方法

2.2.2.1　被试

在清华大学内招募 34 名有偿志愿者（14 名男性，平均年龄为 21.8 ± 3.07）。其中 33 位被试为右利手，1 名为左利手。所有被试视力或者矫正视力正常。实验前被试阅读知情同意书并且签名表示同意。1 名被试由于正确率在随机水平而将其排除。

2.2.2.2　研究材料

实验中采用三个简单的几何图形（正三角形，正方形和圆，每个图形的视角均为 3.7º × 3.7º）与不同道德含义的标签进行匹配。在中央注视点(0.8º × 0.8º)上方呈现图形，下方呈现文字（3.6º × 1.6º）。文字标签分别为（“坏人”、“好人”和“常人”）。图形与标签的匹配关系在被试间进行平衡。注视点与图形和文字中心的距离为 3.5º。被试的任务是判断成对出现的图形与标签是否符合先前学习的过关系（例如，圆形是否是坏人？）

实验在 windows 电脑上完成，刺激呈现的显示器分辨率为 1024×768，刷新频率为 100Hz。实验使用 E-prime 2.0 呈现刺激和收集反应时。

2.2.2.3　程序

被试单个来到实验室参加实验。实验分成两个阶段。首先是学习阶段，让被试学习几何图形（三角形、正方形和圆形）与不同人物标签（好人、坏人和常人）之间的对应关系。例如，我们告诉某个被试，“坏人是圆形；好人是三角形，常人则是由正方形来表示。”学习阶段不给被试呈现几何图形，需要被试自己想像图形与人物之间的关系。当被试表示自己已经记住人物与图形之间的关系之后，开始匹配任务的练习。一般而言，被试能够在 1 分钟以内记住图形与人物之间的联结。

学习之后，进行匹配任务阶段。被试需要判断"图形－人物"的匹配是否正确。每个试次中，首先呈现 500ms 的中央注视点，然后呈现一对"图形－人物"刺激对 100ms。"图形－人物"可能符合指导语中的组合，也有可能不符合。"图形－人物"之后是 1100ms 空屏。被试需要在空屏阶段进行按键，如果符合指导语中的配对，就按 m 键（另一半被试按 n 键）；如果不符合，就按 n（另一半被试按 m 键）。要求被试按键时又快又准（见图 2.1）。被试按键之后，是 500ms 的反馈屏，鼓励被试进行学习。反馈与被试的反应正确程度相一致，如果没有探测到按键，则提示"反应太慢"，让被试加快反应速度。每个 block 结束后，屏幕上会反馈其在整个 block 的总体正确率。

在正式实验之前，每个被试完成一定数目的练习试次，其中匹配与不匹配的试次各一半。被试达的总体正确率达到 60%以后，开始正式实验。

正式实验共 6 个 block，每个 block 包括 60 个试次。因此，每个实验条件均有 60 个试次（坏人－匹配，坏人－不匹配；好人－匹配，好人－不匹配；常人－匹配，常人－不匹配）。

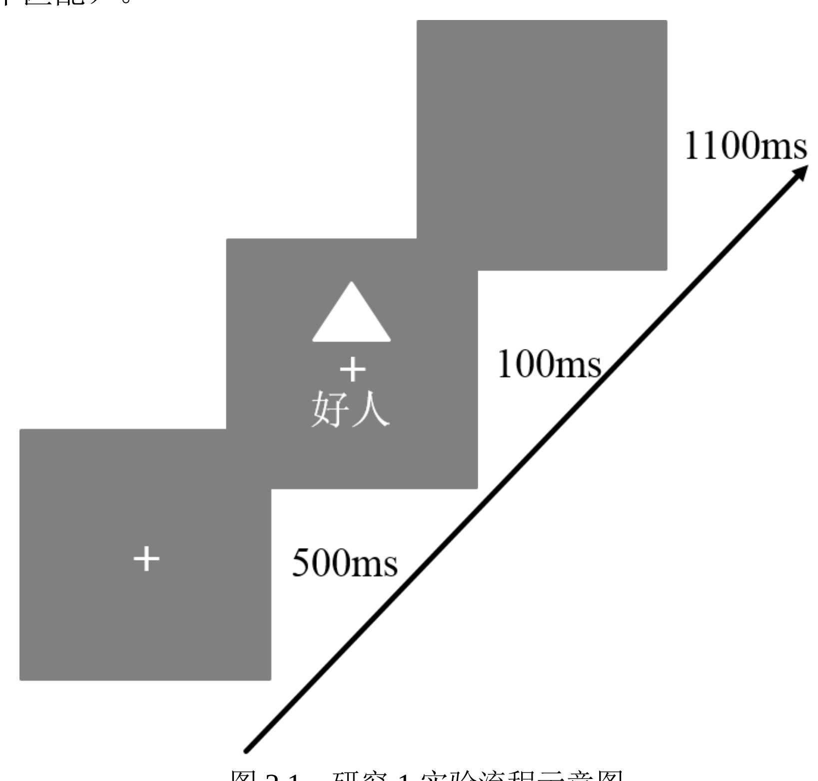


图 2.1　研究 1 实验流程示意图

2.2.2.4 数据分析

初步数据分析表明，被试在反应时和正确率上没有权衡。我们采用信号检测论的方法分析被试的正确率，每种条件下匹配与不匹配条件的数据结果起来计算信

号检测论参数 d'这个指标反映了被试对每种条件下的辨别力，控制了反应倾向的影响。采用重复测量的方差分析对反应时和 d'数据数据进行统计。在假设检验方面，本文报告了效应量和效应量的 90%置信区间；多重比较的情况下采用 Bonferroni 校正。具体而言，根据前人研究的推荐(Bakeman, 2005; Lakens, 2013)，我们报告了 ANOVA 效应量指标 generalized eta square ($\eta_G{}^2$)。

2.2.2　结果

反应时小于 200ms 的正确反应试次未纳入分析，数据少于整体试次的 1%。表 2.1 中显示了研究 1 中各条件下的反应时和正确率。

表 2.1　研究 1 中各条件下的反应时与正确率（括号中为标准差）

Condition	Moral association	Accuracy
Match	Bad person	0.77 (0.15)
	Neutral person	0.82 (0.15)
	Good Person	0.87 (0.13)
Mismatch	Bad person	0.80 (0.14)
	Neutral person	0.78 (0.17)
	Good Person	0.77 (0.13)

对 d'进行重复测量方差分析，以道德效价（好人、坏人和常人）为被试内变量。结果显示道德效价的主效应显著，$F(2,64) = 3.88$, $p = 0.026$, $\eta_p{}^2 = 0.108$, 90% CI [0.01 0.22], $\eta_G{}^2 = 0.029$。采用 Bonferroni 校正的事后比较（post-hoc）表明，好人条件下的 d'显著大于坏人条件($p = 0.01$, Cohen's $d_z = 0.48$)，但是在好人与常人之间以及常人与坏人之间没有显著差异（$ps > 0.15$）（见图 2.2 B）。

对反应时进行重复测量方差分析，以道德效价和匹配条件（匹配 vs.不匹配）为两个被试内变量。结果显示了道德效应的主效应显著，$F(2,64) = 6.19$，$p = 0.003$，$\eta_p{}^2 = 0.162$，90% CI [0.02 0.31]，$\eta_G{}^2 = 0.025$；匹配条件主效应显著，$F(2,64) = 109.1$, $p < 0.0001$, $\eta_p{}^2 = 0.773$, 90% CI [0.68 0.82], $\eta_G{}^2 = 0.13$；交互作用显著，$F(2,64) = 5.84$, $p = 0.005$, $\eta_p{}^2 = 0.154$, 90% CI [0.02 0.30], $\eta_G{}^2 = 0.02$。随后，对匹配和非匹配条件下进行单独的分析。

在匹配条件下，道德效价的主效应显著，$F(2,64) = 6.76$, $p = 0.002$, $\eta_p{}^2 = 0.174$, 90% CI [0.04 0.29], $\eta_G{}^2 = 0.07$。事后检验表明（Bonferroni 校正后的 $p < 0.0167$），

好人条件(M = 703, SD = 78.0)下的反应时显著快于坏人条件(M = 753, SD = 78.4)，$t(32) = 3.54$, $p = 0.001$, Cohen's $d_z = 0.612$，但是与常人条件(M = 735, SD = 79.4)无显著差异，$t(32) = 1.93$, $p = 0.062$。坏人条件与常人条件下也没有显著差异，$t(32) = 2.18$, $p = 0.077$（见图 2.2 A）。

在不匹配条件下，道德效应的主效应不显著，$F(2, 64) = 0.63$, $p = 0.54$。

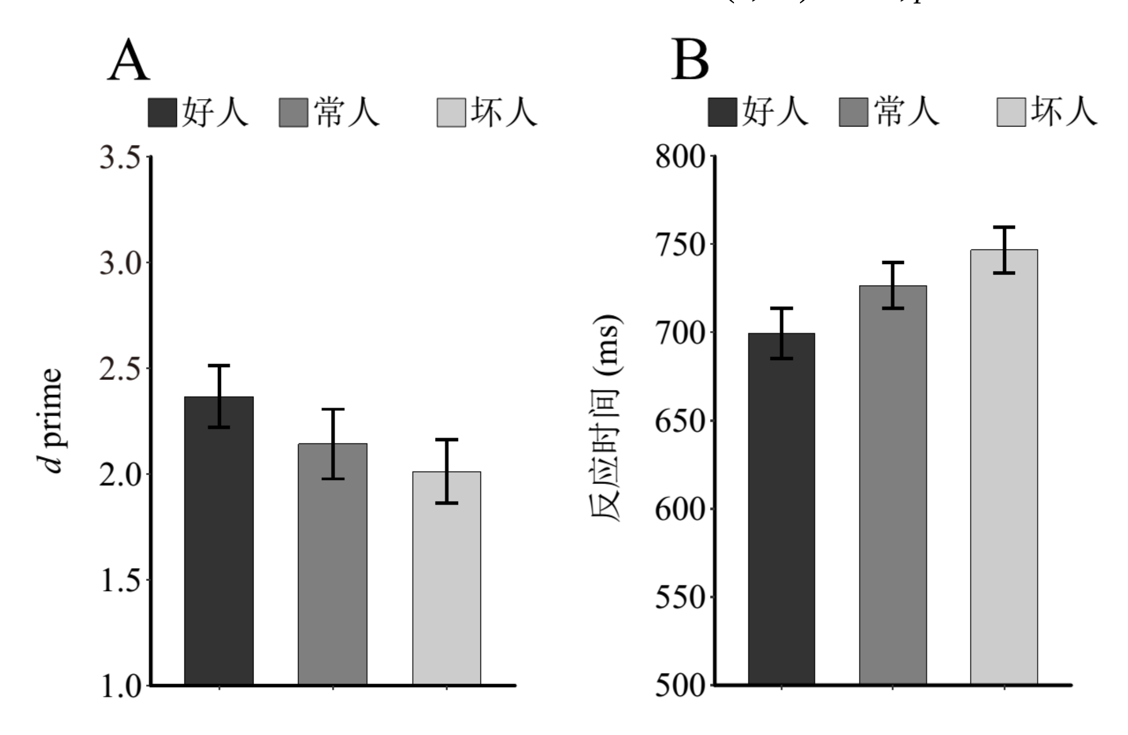


图 2.2　研究 1 中不同条件下的 *d'*（A）及反应时间（B）的均值，误差线为正负一个标准误。

2.2.4　讨论

研究 1 的结果初步表明，道德信息确实会调节知觉加工。具体而言，与消极的道德信息相比，积极的道德信息会促进知觉加工，这个结果在反应时和 d'上均体现出来。这个结果与先前研究中对道德信息的影响一致(Fouragnan et al., 2013)，但是我们知觉任务表明，这种影响可以发生在比决策等高级认知功能更早期的知觉阶段。

有意思的是，我们的研究中，并未发现消极的道德信息的得到优先加工(E. Anderson et al., 2011)。这种结果的差异可能是由于任务所致：在 E. Anderson et al. (2011)的任务中，他们使用的是探测任务，被试需要探测出消极刺激的存在，因此消极刺激本身就是目标；而在我们的任务中，被试的任务是判断图形—人物的匹配关系，此时消极刺激可能作为干扰，其对注意的吸引可能减慢了对任务的反应。这

与消极情绪面孔对注意的捕获类似(Eastwood, Smilek, & Merikle, 2003)。

研究 1 的结果可能是由于道德信息的影响，也有可能是由于人物标签熟悉性的影响。“好人”和“坏人”比“常人”使用得更加频繁，在熟悉程度上更高。为了排除熟悉性的影响，我们进行了研究 2。在研究 2 中，我们先对与“好人”、“坏人”和“常人”相似的词在熟悉程度、具体性和使用频率上进行调查，然后，选择在熟悉程度上相匹配的词，重复研究 1 的程序。

2.3　研究 2：排除熟悉性的影响

2.3.1　引言

研究 1 已经初步表明，积极的道德信息确实会影响知觉匹配的加工。但是这可能是由于词汇本身熟悉性的影响，为了控制熟悉性的影响，我们采用熟悉性上没有差异的词汇作为实验刺激。研究 2 采用与研究 1 相同的方法，主要的区别将词汇换成熟悉性上没有差别的词。为了对语汇的熟悉性进行评定，在研究 2 之前，我们进行了词汇的评定任务。

2.3.2　方法

2.3.2.1　实验材料

为了对词汇的熟悉程度进行评定，在正式实验前，招募 35 人与正式实验同质的独立样本（13 名男性，年龄为 20.6 ± 3.11），分别对与“好人”、“坏人”和“常人”以及其同义词进行评估。评定中使用的词语包括：好人、坏人、常人、善人、恶人、凡人、君子、流氓。

被试需要从如下几个维度对每个词所代表的人进行评分：熟悉性、使用频率和具体性。正式实验中，选择“善人”、“恶人”和“常人”作为实验的标签。评定结果表明，这三个词在熟悉性和具体性上无差异，但是常人的使用频率要显著高于善人和恶人（图 2.3）。

2.3.2.2　被试

39 名课程被试参加了研究 2，所有被试均是右利手，视力或者校正视力正常。实验前，被试阅读并签署了《知情同意书》。与实验 1 相似，12 名被试在正式实验数据中，整体正确率低于 60%的数据被删除。剩下的 27 名被试中，10 名为男性，年龄 20.07 ± 2.3。

2.3.2.3　实验程序

实验程序与研究 1 相同。

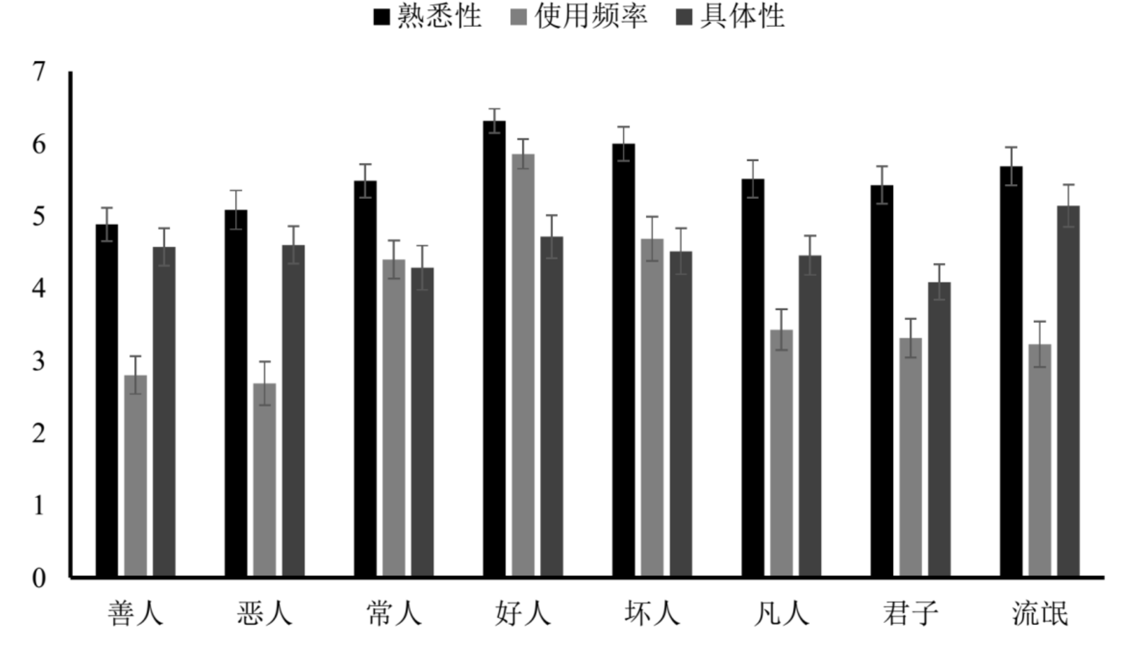


图 2.3　各个词汇标签在熟悉性、使用频率和具体性上的得分。

2.3.3　结果

与研究 1 相同，我们对反应时和 *d'*进行了重复测量方差分析，道德效价和匹配程度为两个被试内因素。

在 *d'*上，发现道德效价的主效应：$F(2, 52) = 8.78, p = 0.001, \eta_p^2 = 0.25$, 90% CI [0.08 0.38]，$\eta_G^2 = 0.06$。对匹配的试次进行各条件的配对样本 *t* 检验（Bonferroni 校正的阈值为 $p < 0.0167$），结果显示，善人条件(2.24 ± 0.76)下显著高于常人条件(1.78 ± 0.73)，$t(26) = 4.48, p < 0.0001$, Cohen's $d_z = 0.86$，同时边缘显著地高于恶人条件(1.94 ± 0.73), $t(26) = 2.17, p = 0.039$, Cohen's $d_z = 0.42$。但是在常人和恶人条件下，没有显著的差异，$t(26) = 1.83, p = 0.079$, Cohen's $d_z = 0.35$。

对反应时数据的重复测量方差分析表明，道德效价的主效应显著：$F(2,52) = 19.36, p < 0.0001, \eta_p^2 = 0.43$, 90% CI [0.24 0.54], $\eta_G^2 = 0.043$；匹配程度的主效应显著 $F(1,26) = 10.4, p < 0.0001, \eta_p^2 = 0.779$, 90% CI [0.07 0.47], $\eta_G^2 = 0.12$。道德效价与匹配程度之间交互作用显著：$F(2,52) = 5.76, p = 0.006, \eta_p^2 = 0.18$, 90% CI [0.03 0.31], $\eta_G^2 = 0.019$。Bonferroni 校正后的配对样本 *t* 检验表明，善人条件下的反应时(671.3 ± 74.1)显著快于常人(671.3 ± 74.1)（$t(26) = 5.17, p < 0.0001$, Cohen's $d_z = 0.99$）和恶人条件(716.7 ± 71.7)（$t(26) = 3.8, p = 0.001$, Cohen's dz = 0.73）；但是恶人与常人条件没有显著差异 $t(26) = -0.94, p = 0.358$。

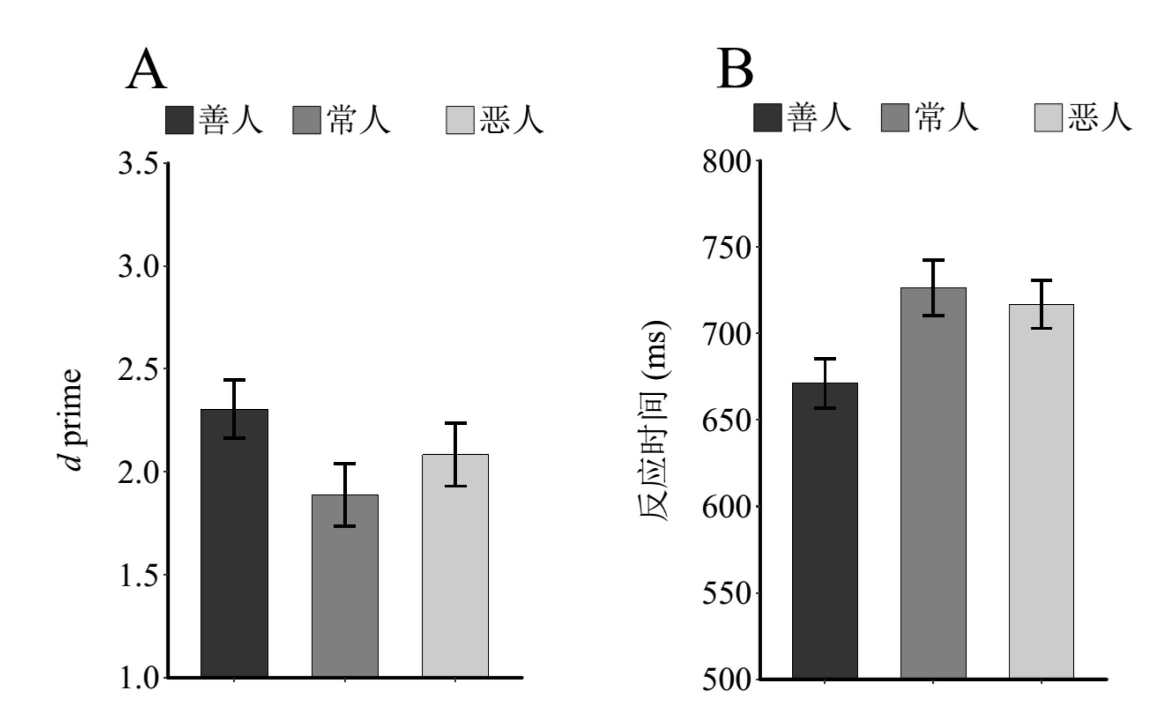


图 2.4　研究 2 中不同条件下的 *d'*（A）及反应时间（B）的均值，误差线为正负一个标准误。

跨实验分析　为了进一步检验道德信息调节效应的稳定性，我们联合两个实验的结果，以实验（1 vs. 2）作为被试间变量，以道德效价和匹配程度作为被试内变量进行重复测量方差分析。对反应时进行了 2 (实验：1 vs. 2) × 2 (匹配条件：匹配 vs. 不匹配) × 3 (道德效价：积极、中性、消极) 的重复测量方差分析，结果表明在实验与其他变量之间的交互作用均不显著($ps > 0.165$)。对 *d'* 进行了 2 (实验：1 vs. 2) × 3 (道德效价：积极、中性、消极) 重复测量方差分析，结果再次发现，在实验与道德效价之间没有交互作用 ($F = 1.7, p = 0.95$)。结果表明，两个实验作为组间变量并未与道德或者匹配程度产生交互作用，说明道德信息的调节作用具有跨实验的一致性。

2.3.4　讨论

研究 2 的主要目的在于检验研究 1 中效应是否是由于熟悉性因素的影响。在本实验中，我们使用了熟悉性和具体上均无显著差异的词汇作为人物标签，其中“常人”的使用频率要显著高于“善人”和“恶人”。但是本实验的结果表明，积极的道德信息（善人）在正确率（使用 *d'* 作为指标）和反应时上均要优于中性和消极道德的情况。

本实验的结果表明，对词汇的熟悉性无法解释研究 1 中的效应。跨实验的分

析表明，两个实验之间没有交互作用，进一步表明词汇的熟悉性不影响道德信息在知觉加工中的优势作用。

为进一步检验积极道德信息对知觉的促进作用，我们进行了研究 3。研究 3 中将人物标签先呈现，然后再呈现图形。先呈现标签时被试可以形成对图形的预期，从而会加快其反应(Sui, Sun, Peng, & Humphreys, 2014)。 但是，由于道德信息的突显性，我们预期，先呈现标签也不会影响到道德信息的调节作用。

2.4　研究 3：排除文字对知觉加工优势的影响

2.4.1　引言

研究 1 和 2 表明，道德信息，尤其是积极的道德信息能够促进在知觉匹配任务中的表现。但是由于研究 1 与 2 中，被试所看到的刺激均为图形与文字的组合，因此有可能被试对道德上积极性的快速加工，可能是对文字本身的加工在而非对于学习到图形－文字联结的加工。因此在研究 3 中，先呈现文字标签，同去呈现图形，被试仅需要对图形进行反应。一方面，这是对研究 1 和 2 的概念重复，消除了文字的影响；另一方面，由于先呈现文字标签时，人们往往会形成对随后刺激的预期从而有更快的知觉加工，研究 3 可以考察这种预期是否能够解释道德信息对知觉的影响。

2.4.2　方法

2.4.2.1　被试

35 名大学生被试参加了本次实验，其中男性 18 名，年龄 21.6 ± 2.8 岁。其中 24 名被试参加过研究 1（9 男，年龄为 21.9 ± 2.9）。在 35 名被试之中，2 名被试的数据被删除，其中一名被试实验后反馈其右手中指有伤，而该手指是用来按键反应的；另一名被试则是由于正式实验中正确率未达到 60%。数据分析过程与研究 1 相同。

2.4.2.2　实验材料与程序

研究 3 的程序与研究 1 相比有如下变化：第一，人物标签与图形序列呈现，人物标签呈现 50ms，图形呈现 50ms；第二，人物标签之后呈现一个打乱图片 200ms，用来打断对人物标签的进一步加工；第三，为了避免正确率上的天花板效应，目标图形呈现在噪音背景中，噪音背景是由 3/4 灰度和 1/4 白色的图片打乱成为 2 像素×2 像素的像素点图。噪音背景的混乱程度通过预实验测试有效的。因此，在实验 2 中，每个试次开始呈现 500ms 的注视点，然后呈现人物标签（“好人”、“坏人”或“常人”）50ms，然后呈现掩蔽刺激 200ms。掩蔽刺激之后再呈现目标图形 50ms，接下来被试有 1100ms 的时间窗口进行按键反应。与研究 1 相同，500ms 呈现对被

试进行反馈，每个 block 结束后呈现反馈整体的正确率和平均反应时。下一个试次间隔 700~1100ms 后开始（见图 2.5）。

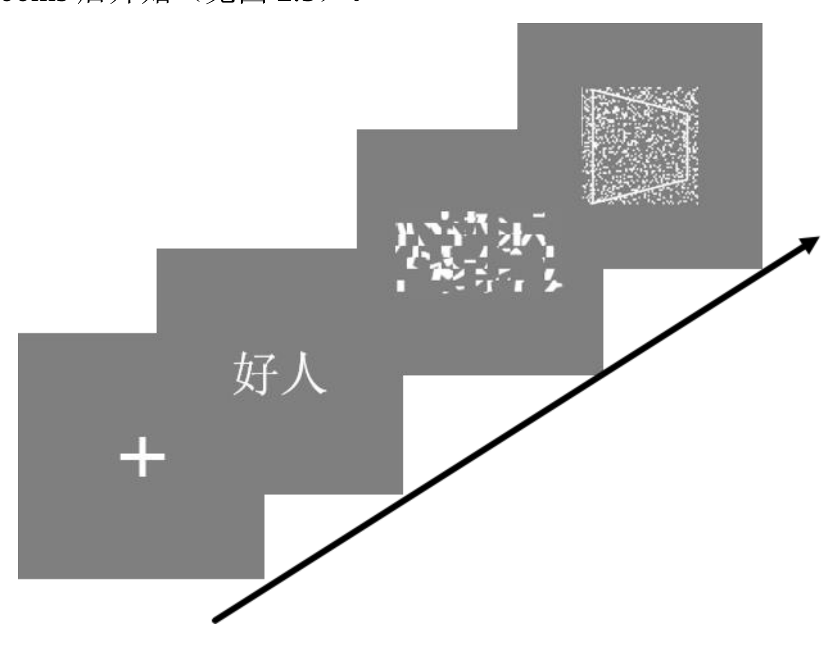


图 2.5　研究 3 的刺激与实验流程图

2.4.3　结果

2.4.3.1　*d'* 与 RT 的结果

与研究 1 相同，我们对反应时和 *d'*进行了重复测量的方差分析。

对 *d'*的重复测量方差分析表明，道德效价的主效应显著：$F(2, 64) = 13.55$, $p < 0.0001$, $\eta_p^2 = 0.297$, 90% CI [0.14 0.42], $\eta_G^2 = 0.06720$。与实验 1 相同，好人条件(M = 2.86, SD = 1)下的 *d'*大于好人(M = 2.48 , SD = 0.79, $t(32) = 3.97$, $p < 0.0001$, Cohen's d_z =0.69)和常人条件 (M = 2.28 , SD = 0.91, $t(32) = 4.5$, $p < 0.0001$, Cohen's $d_z = 0.78$)；坏人与常人条件下没有显著差异，$t(32) = 1.77$, $p = 0.086$, Cohen's $d_z = 0.31$　见图 2.6）。

反应的分析也得到了与研究 1 相似的结果。道德效价的主效应显著 $F(2,64) = 6.02$ $p = 0.004$, $\eta_p^2 = 0.158$, 90% CI [0.03 0.28], $\eta_G^2 = 0.02056$；匹配程度的主效应显著($F(1,32) = 129.7$, $p < 0.001$, $\eta_p^2 = 0.802$, 90% CI [0.68 0.86], $\eta_G^2 = 0.11029$；两者的交互作用显著 $F(2, 64) = 10.16$, $p < 0.0001$, $\eta_p^2 = 0.24$, 90% CI [0.09 0.36], $\eta_G^2 = 0.0187$。分别对匹配和不匹配条件下的席次进行分析，在匹配条件下，道德效价的主效应显著：$F(2,64) = 10.55$, $p < 0.0001$, $\eta_p^2 = 0.248$, 90% CI [0.1 0.37], $\eta_G^2 = 0.069$。Bonferroni 校正后的 *t* 检验表明，好人条件下(M = 550, SD = 55)的反应时显著快于常人条件(M = 583, SD = 64)，$t(32) = 3.76$, $p = 0.001$, Cohen's $d_z = 0.65$；好人条件下的反应也

快于坏人条件(M = 585, SD = 59)，$t(32) = 4$, $p < 0.0001$, Cohen's $d_z = 0.697$；坏人与常人没有显著差异，$t(32) = 0.27$, $p = 0.79$, Cohen's $d_z = 0.0478$。不匹配试次中，道德效价主效应不显著。$F(2,64) = 0.157$, $p = 0.855$。

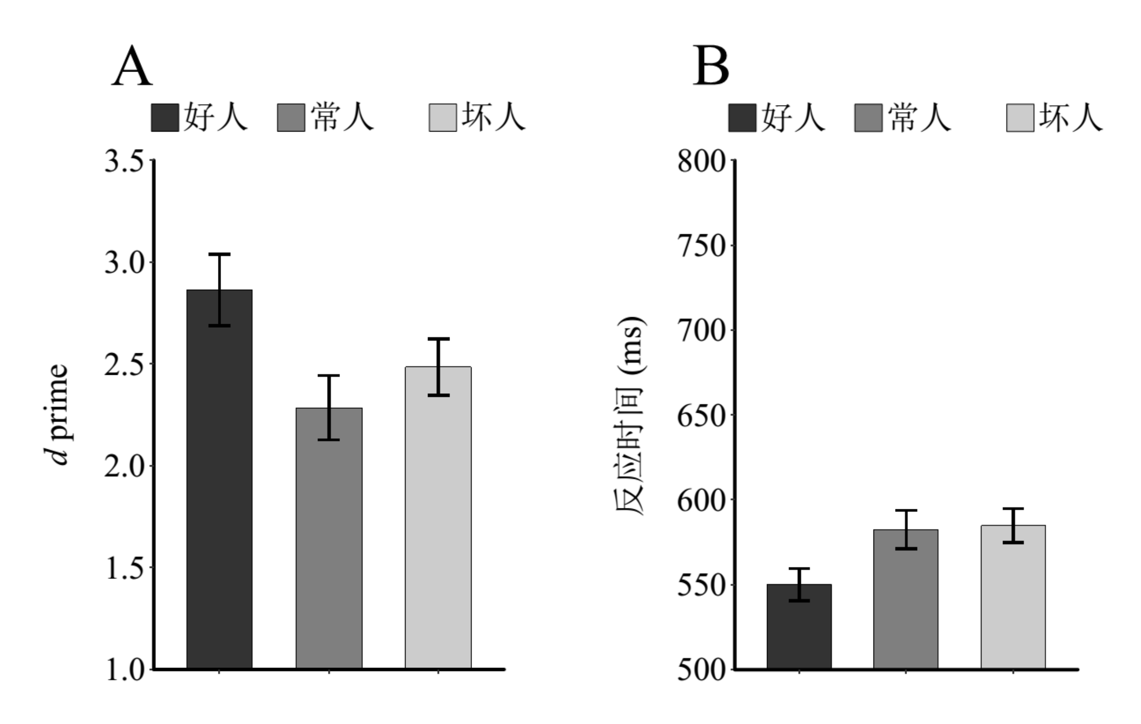


图 2.6　研究 3 中不同条件下的 *d'*（A）及反应时间（B）的均值，误差线为正负一个标准误。

2.4.3.2 跨实验分析

为检验道德信息对知觉调节作用在个体身上的稳定性，对同时参加过两个实验的被试进行跨实验的分析。以道德效价（好、坏和常）和实验（1 vs. 3）作为被试内变量。重复测量方差分析表明，道德效价主效应显著，$F(2,46) = 10.97$, $p < 0.0001$, $\eta_p^2 = 0.32$，实验主效应显著，$F(1,23) = 220.8$, $p < 0.0001$, $\eta_p^2 = 0.90$；两者之间的交互作用不显著 $F(1.6,36.7) = 2.95$, $p = 0.076$。事后比较发现，好人条件(614.8 ± 12.3)下的的反应时显著适快于坏人(656 ± 11, $p = 0.006$)和常人条件(658 ± 12.5, $p = 0.002$).。

2.4.4　讨论

研究 3 的结果重复研究 1 中出现的结果模式：积极的道德信息促进了知觉匹配的加工。

实验 1、2 和 3 非常一致地发现，积极道德信息对知觉的促进作用。这个结果表明，即使是在实验室中即时学习到的道德意义，也能够对知觉产生影响。由于本

研究中不仅使用了三个简单的词汇标签，因此不可能是由于语义启动造成的影响(Firestone & Scholl, 2014)。

此外，本研究中，使用社会联结学习的范式，刺激相对简单明确，而且刺激的联结关系在被试间进行了平衡，排除了被试自发的特质推断的影响。

本研究中所发现的结果，与先前研究中所发现的“欺骗检测理论（cheater-detection theory）不一致(Cosmides, 1989; Cosmides & Tooby, 2000)。欺骗检测理论认为，人们会对潜在的欺骗者更加敏感。然而， 我们的研究结果表明，积极的刺激更具有加工的优势。可能的原因是，消极信息可以自动地吸引注意，但是当消极信息不是与任务相关时，它会阻碍对任务的加工。

积极的道德信息为什么能力在知觉加工中具有优势？一个可能的机制是由于对于积极的道德信息加工中，有一种内在的道德自我参照，即人们内隐地将积极的道德信息作为自我的部分。另一个可能的机制是道德上积极的信息对个体是有利的，出于趋利避害的本能，人们会更倾向于趋近道德高尚的人。

如果积极的道德信息在知觉中的优势是由于内在的自我参照，则这种优势效应将受到自我相关性的调节：在自我相关的情况下，我们会对积极的道德信息具有加工优势，而在与自我不相关的情况下，则这种优势效应消失。与内在道德自我参照假设不同的，如果积极道德信息中是由于积极道德效应的潜在利益，则这种积极道德的优势不受到自我的调节。

为了检验这两种不同解释，我们进行了研究 4，加入了自我相关性的变量，探讨道德自我是道德信息加工内在参照理论。

第三章　道德自我的知觉加工优势

3.1　本章引言

如前所述，道德信息是社会知觉的基本维度，是对他人印象形成中最重要的维度之一。在前人的基础之上，研究 1、2 和 3 采用社会联结学习的范式，发现道德品质、尤其是积极道德品质，能够促进人们的知觉。这表明道德信息能够通过自上而下的调节，在认知加工的早期阶段起作用。

但是道德信息在知觉中加工的优势的原因是什么呢？一种解释从道德自我作为内在参照出发，认为人们在进行信息加工时，可能使用道德自我作为锚定点进行参照，积极的道德品质（好人）与道德自我更加接近，而中性和消极的道德品质与自我更远，因此对积极的道德品质加工更快。

这种道德自我内隐参照的解释来自于自我相关信息在知觉加工中的优势：社会心理学的研究发现，人会觉得自己在各个方面高于平均水平(Alicke & Govorun, 2005)；而认知心理学的研究发现，人们在记忆、注意和知觉方面，均对自己有相关的信息有优势(Sui & Humphreys, 2015a)。最近的研究表明，人们对于自己的道德形象认识中具有自我增强的效应，而且这种效应可能会比其他方面更强(Tappin & McKay, 2016)。这种增强效应已经在记忆方面得到了实证的支持(Kouchaki & Gino, 2016; Sedikides et al., 2016)。

但是另一种解释认为，对积极道德信息的加工优势，仅仅是由于积极道德信息可能反映了潜在的奖赏，即具有积极道德品质的人可能是更好的合作者，因此更有可能给我们带来奖赏。而奖赏信息能够在知觉中得到快速加工(B. A. Anderson et al., 2011)。

为了检验这两种可能的解释，我们进行了研究 4、5 和 6。在研究 4 中，我们外显地将让被试学习自我和他人在道德上积极、消极和中性的三个方面，将不同的方面与不同的几何图形联系起来，然后进行知觉匹配的任务。研究 5 中，我们让被试仅学习自我和他人与几何图形的联结，但是将道德品质信息呈现在图形之中作为无关刺激，从而考查道德品质信息作为任务无关刺激时，对知觉决策的影响。研究 6 与研究 5 思路相似，但是将自我相关性与道德信息的位置互换，让被试学习道德品质信息，将自我和他人呈现在图形之中作为无关刺激，从而探索自我相关性对道德品质信息在知觉中的内隐调节效应。

3.2　研究 4：道德自我的知觉加工优势

3.2.1　引言

为了直接验证道德信息加工优势中的内在道德自我参照解释与潜在奖赏的解释，研究 4 在研究 1 的基础之上，加入新的变量：自我相关性。在本实验中，让被试将自己在道德意义上积极、中性和消极的方面与他人在道德意义上积极、中性和消极的方面与几何图形进行联结，再进行知觉匹配任务，从而能够直接考虑道德自我条件下的知觉加工特点。

从理论上讲，本研究可以验证三个不同的理论假设。第一，道德效价的主效应，如果主效应显著，表明道德信息的对知觉的加工中，主要是由于潜在的奖赏带来的优势，验证了潜在奖赏假设。

第二，自我相关性的主效应，即自我条件与他人条件相比的差异。如果自我主效应存在，表明不同方面的自我对知觉的影响相同，这将表明自我相关性具有非常强的优势，引入自我相关性之后，道德信息的优势不复存在，从而与先前关于自我的研究一致。

第三，自我相关性与道德效价的交互作用。如果交互作用显著，则表明，自我与道德的效应并非单个因素所影响，两者是相互影响的变量。更进一步而言，如果道德效价因素仅在自我条件下起作用，则表明道德品质信息可能受到内在道德自我的调节，即验证了内隐道德自我参照的假设；如果道德效价因素仅在他人条件下起作用，则表明道德信息的加工优势可能受到奖赏的调节，但与自我的关系不明确。

我们预期，道德效价与自我相关性将产生交互作用，而且与内隐道德自我参照假设一致。

3.2.2　方法

3.2.2.1　被试

27 名大学生参加了本次实验，其中男性 15 名，年龄为 19 到 26 岁，21.59 ± 2.05。所有的被试均是右利手，视力或者校正视力正确。实验前，所有被试签署《知情同意书》。3 名被试的数据被排除出分析，其中一名是由于正式实验中正确率过低，另外两名则是由于实验中途退出。最终有效数据为 24 名被试，平均年龄为 21.79 ± 2.08 岁。

3.2.2.2 实验设计

本实验为 2× 3× 2 完全被试内实验设计。第一个变量为自我相关性，包括自我和陌生人两个水平；第二个自变量为道德效价，包括积极、消极和中性的道德水平；第三个自变量为匹配程度，包括匹配和不匹配。

3.2.2.3 实验材料

本实验的实验材料与研究 1 的参数相同，但是增加了三个几何图形。因此，几何图形包括三角形、正方形、圆、梯形、梯形和正五边形，人物标签也包括 6 个：好我、凡我、坏我、好人、凡人和坏人。

为了匹配不同人物标签之前具体性，实验前让被试从一系列男性和女性的中文名字中，选择出一个他/她认为最陌生的同性的名字。开始练习前，告诉被试每个人物标签的意义，比如“好我”代表的是自己道德上高尚的一面，可以想像自己做了某件道德高尚事件时刻的自己；“坏我”代表的是自己道德败坏的一面，可以想像自己做了某件道德败坏事情时刻的自己；“凡我”代表自己普通的一面，可以想像自己做了与道德无关事情时刻的自己。“好人”、“坏人”和“凡人”则是指的被试自己所选择名字的那个陌生人，也指的他人三个不同的侧面。向被试解释这六个标签的意义，直到被试能够准确理解这六个标签所代表的意义为止。

2.2.2.4 实验程序

研究 4 的实验程序与研究 1 相同，共 6 个 block，每个 block 有 120 个试次。

3.2.3 结果

研究 4 结果的分析方法与研究 1 相同。

对于 d’的重复测量方差分析表明，道德效价与自我相关性的交互作用显著：$F(2,46) = 3.63$, $p = 0.034$, $\eta_p^2 = 0.136$, 90% CI [0.01 0.27], $\eta_G^2 = 0.038$。道德效价的主效应显著：$F(2,46) = 5.03$, $p = 0.011$, $\eta_p^2 = 0.18$, 90% CI [0.03 0.32], $\eta_G^2 = 0.036$。但是自我相关性的主效应不显著：$F(2.46) = 1.2$, $p = 0.28$。因此，对自我和陌生人的条件进行分别的分析。在自我条件下，道德效价的主效应显著：$F(2, 46) = 9.44$, $p < 0.0001$, $\eta_p^2 = 0.29$, 90% CI [0.1 0.43], $\eta_G^2 = 0.016$。简单效应分析表明，好我条件下的 d’ (M=2.35 ± 0.8)显著大于坏我(M=1.81 ± 0.46) ($t(23) = 3.04$, $p = 0.006$, Cohen’s $d_z = 0.62$) 和凡我条件 (M = 1.78 ± 0.51) ($t(23) = 3.71$, $p = 0.001$, Cohen’s $d_z = 0.757$)，

但是在坏我和凡我之间没有显著差异，$t(23) = 0.25$, $p = 0.81$。对他人条件下 d’的方差分析表明，道德效价的主效应不显著，$F = 0.009$, $p = 0.99$。

此外，还对不同道德效价水平下，自我与他人在 d’上的差异。结果显示，坏人(2.12 ± 0.64)比坏我(1.81 ± 0.46)的表现更佳 $t(23) = 2.35$, $p = 0.028$, Cohen’s $dz = 0.48$；凡人(2.1 ± 0.82)比凡我(1.8 ± 0.53)表现边缘显著的优势，$t(23) = 1.95$, $p = 0.063$, Cohen’s $dz = 0.398$；但是好人(2.1 ± 0.7)与好我(2.3 ± 0.8)没有显著差异，$t(23) = 1.03$, $p = 0.31$。

对反应时的重复测量方差分析发现，道德效价的主效应显著，$F(2,46) = 16.7$, $p < 0.0001$, $\eta_p^2 = 0.42$；同样显著的有匹配程度的主效应，$F(1,23) = 112.7$，$p < 0.0001$, $\eta_p^2 = 0.83$；但是自我相关性的主效应不显著 $F(1,23) = 0.31$, $p = 0.58$. 三者的交互作用不显著，$F(2,46) = 0.69$, $p = 0.51$。

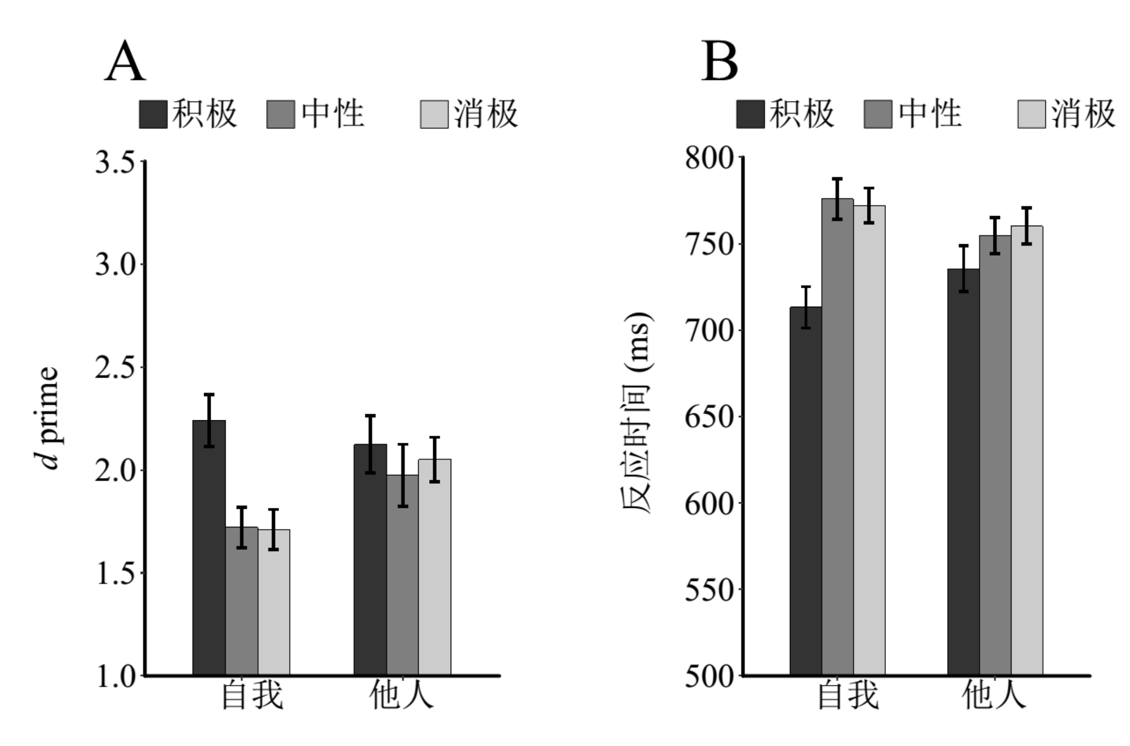


图 3.1　研究 4 中不同条件下的 *d’*（A）及反应时间（B）的均值，误差线为正负一个标准误。

为了进一步分析变量之间的关系，我们仅考虑匹配的试次。重复测量方差分析表明，道德效价与自我相关性的交互作用显著 $F(2.46) = 3.361$, $p = 0.043$, $\eta_p^2 = 0.127$, 90% CI [0 0.26], $\eta_G^2 = 0.025$；道德效价的主效应显著，F(2.46) = 21.25, $p < 0.0001$, $\eta_p^2 = 0.48$, 90% CI [0.28 0.59], $\eta_G^2 = 0.067$。

随后，对自我和他人条件的试次进行单独分析。对自我条件下匹配试次的反应时分析表明，道德效价的主效应显著：$F(2,64) = 26.26$, $p < 0.0001$, η_p^2 =0.53, 90%

CI [0.34 0.63], $\eta_G{}^2 = 0.16$。Bonferroni 校正的配对样本 t 检验表明，好我条件(M=716, SD = 66.0)比坏我条件(M=770, SD = 65.5)的反应时更快，$t(23) = 4.68$, $p < 0.0001$, Cohen's $d_z = 0.95$, 也比凡我条件(M = 781, SD = 64.2)更快，$t(23) = 8.24$, $p < 0.0001$, Cohen's $d_z = 1.68$。但是凡我与坏我之间没有显著差异，$t(23) = -1.22$, $p = 0.23$, Cohen's $d_z = 0.249$。对他人条件下匹配试次的分析表明，道德效价的主效应不显著：$F(2,46) = 0.84$, $p = 0.44$。

此外，对不同道德条件下的自我与他人反应时的比较发现，凡人 (759 ± 58) 比凡我条件(781 ± 64)更快，连到边缘显著：$t(23) = 1.87$, $p = 0.075$, Cohen's $dz = 0.38$。但是在其他条件下自我与他人无显著差异($ps > 0.2$)。

表 3-1. 研究 4 中各种条件下的正确率

条件	身份	道德效价	正确率
匹配	自我	消极	0.76(0.13)
		中性	0.76(0.15)
		积极	0.88(0.13)
	他人	消极	0.82(0.13)
		中性	0.81(0.13)
		积极	0.84(0.14)
不匹配	自我	消极	0.76(0.11)
		中性	0.75(0.12)
		积极	0.73(0.14)
	他人	消极	0.79(0.12)
		中性	0.77(0.14)
		积极	0.77(0.12)

3.2.4　讨论

研究 4 联合了自我相关性与道德效价来进一步探索道德信息调节知觉加工的

机制。实验结果发现，积极道德信息对知觉的促进效应仅存在于自我条件。这个结果模式验证了内隐道德自我参照的假设。

研究四的结果与先前的研究一致，再次验证了自我的积极偏见（positive bias of self），可能是被试认为只有积极方面的我才是真我(Newman, Bloom, et al., 2014)。这些结果与关于道德自我的记忆研究也具有一致性，即被试更倾向于忘记自己做过的不道德行为(Kouchaki & Gino, 2016)。

研究 4 的结果也表明，道德信息在知觉中的加工作用，并不仅仅是由于积极效价的作用，自我起到了关键的作用。

为了进一步验证内隐道德自我参照的假设，我们进一步研究道德与自我的交互作用，研究 5 和 6 分别以道德和自我为无关刺激，测量其对自我和道德加工的影响。

3.3　研究 5：道德效价的内隐调节作用

3.3.1　引言

通过研究 1～3，本研究发现，道德信息在知觉中的具有优势。而研究 4 表明，这种优势的原因可能是由于内隐的道德自我参照加工。由于研究 4 采用的是外显的任务，因此无法直接说明道德自我的内隐参照的特点。此外，研究 4 中被试需要学习 6 个图形－标签对，工作记忆负荷较大。

为了更进一步地探索道德自我的内隐参照假设，本研究拟采用内隐的方式，让自我信息或者道德信息作为与任务无关的刺激呈现，检验其出现是否会对知觉反应起到促进作用。

在研究 5 中，以自我联结任务为外显的任务，以道德刺激作为分心物，以检验在道德信息作为与任务无关的刺激时，是否会影响到被试的任务表现。我们预期，当积极的道德信息与自我信息结合时，会促进知觉加工，而当中性或者消极的信息与自我信息结合时，则不会产生这种优势。

3.3.2　方法

3.3.2.1　被试

31 名大学生参加了本次实验，其中男性 16 名，平均年龄为 19.39 ± 1.23。所有的被试均是右利手，视力或者校正视力正确。实验前，所有被试签署《知情同意书》。3 名被试的数据由于总体的正确率接近随机水平而被排除出分析。最终有效数据为 28 名被试（15 名男性），年龄为 19.46 ± 1.2 岁。

3.3.2.2　实验设计

本实验为 2× 3× 2 完全被试内实验设计。第一个变量为自我相关性，包括自我和陌生人两个水平；第二个自变量为道德效价，包括积极、消极和中性的道德水平；第三个自变量为匹配程度，包括匹配和不匹配。

3.3.2.3　实验材料

本实验的实验材料与研究 4 的参数相同，但是仅使用两个几何图形：正方形

和圆形。人物标签仅两个：自我和他人。

与研究 4 相同，实验前让被试从一系列男性和女性的中文名字中，选择出一个他/她认为最陌生的同性的名字。开始练习前，告诉被试每个人物标签的意义，比如“自己”代表的被试自己；而“他人”则代表的是刚刚选择的这个名字所代表的人。告诉被试自我与他人与几何图形的对应关系，然后被试开始练习。

练习时，每个图形的中间均可能会有一个词：好人、坏人或者常人。要求被试忽略图形中的文字，仅观察图形与正文标签之间的匹配关系。

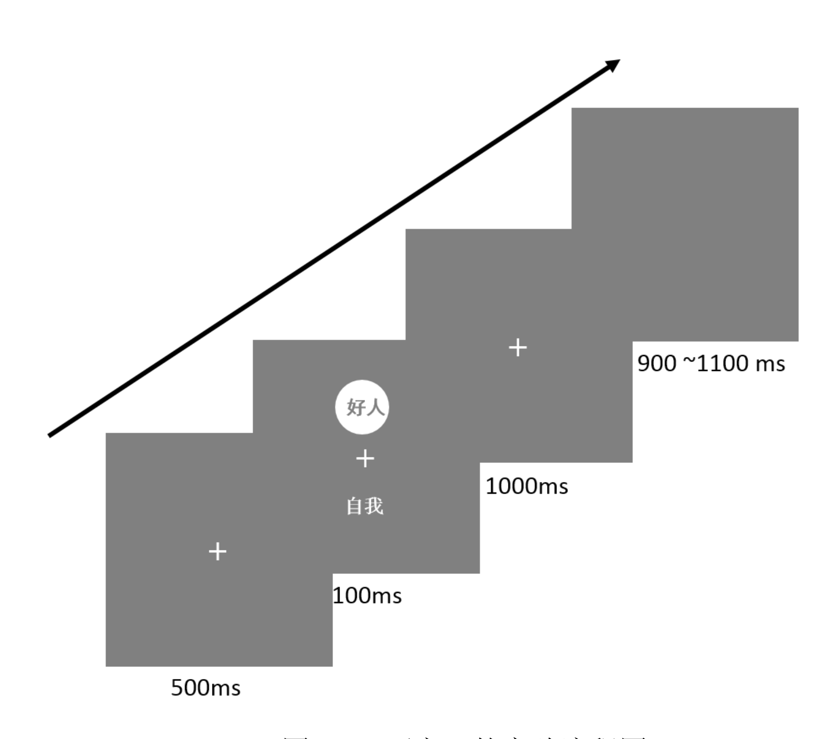


图 3.2　研究 5 的实验流程图

3.3.2.4　实验程序

研究 5 实验程序与研究 4 相同，共 6 个 block，每个 block 有 60 个试次。

3.3.3　结果

研究 5 结果的分析方法与研究 4 相同。

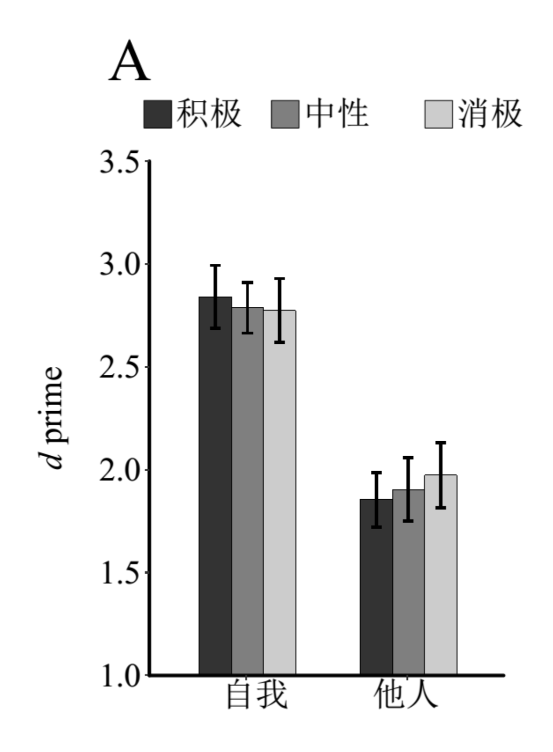
对于 *d'*的重复测量方差分析表明，自我相关性的主效应显著：$F(1, 27) = 49.67$, $p < 0.0001$, $\eta_g^2 = 0.2543$，自我联结条件下的 *d'* (2.798 ± 0.754) 比他人联结条件下 (1.906 ± 0.783)更高, $t(27) = 7.048$, $p < 0.0001$, *Cohen's* $d_z = 1.3319$, 95% CI [0.92 1.821]。道德效价的主效应不显著 $F(2, 54) = 0.115$, $p = 0.8915$, $\eta_g^2 = 0.0003$；道德

效价与自我参照的交互作用也不显著 $F(2, 54) = 0.658$, $p = 0.5222$, $\eta_g^2 = 0.0022$。

对反应时的重复测量方差分析发现，自我参照的主效应显著，$F(1, 27) = 103.946$, $p < 0.0001$, $\eta_g^2 = 0.2943$；道德效价与自我参照的交互作用显著，$F(2, 54) = 3.957$, $p = 0.0249$, $\eta_g^2 = 0.006$；道德效价的主效应不显著 $F(2, 54) = 1.404$, $p = 0.2544$, $\eta_g^2 = 0.0033$。

为了进一步分析变量之间的关系，我们仅考虑匹配的试次。对于自我联结的条件，有显著的道德效价的主效应，$F(2, 54) = 7.245$, $p = 0.0016$, $\eta_g^2 = 0.0173$. 配对样本 t 检验显示道德自我条件下 (634 ± 70)的反应时间比不道德自我条件下(655 ± 67)的反应时更快, $t(27) = -4.228$, $p = 0.00024$, $Cohen's\ d_z = -0.799$, 95% CI [-1.172 -0.429]，也快于常我的条件(650 ± 66): $t(27) = -2.67$, $p = 0.01269$, $Cohen's\ d_z = -0.5045$, 95% CI [-0.917 -0.077]. 常我与不道德自我条件下的反应时间没有显著差异，$t(27) = 0.726$, $p = 0.47379$, $Cohen's\ d_z = 0.1373$, 95% CI [-0.253 0.525]。在他人条件下，则没有道德效价的主效应，$F(2, 54) = 0.079$, $p = 0.9244$, $\eta_g^2 = 0.0004$。

我们也检验了每种条件下的自我参照效应。配对样本 t 检验显示，道德自我条件下(634 ± 70) 要显著快于道德他人的条件(732 ± 63) : $t(27) = -11.324$, $p < 0.0001$, $Cohen's\ d_z = -2.14$, 95% CI [-2.72 -1.573]；平均自我条件下反应时间(650 ± 66) 快于平均他人的反应时间 (730 ± 70), $t(27) = -7.805$, $p < 0.0001$, $Cohen's\ d_z = -1.4751$, 95% CI [-2.051 -0.977]；不道德自我条件下的反应时(655 ± 67) 短于不道德他人(730 ± 64): $t(27) = -7.384$, $p < 0.0001$, $Cohen's\ d_z = -1.3955$, 95% CI [-1.953 -0.753]。



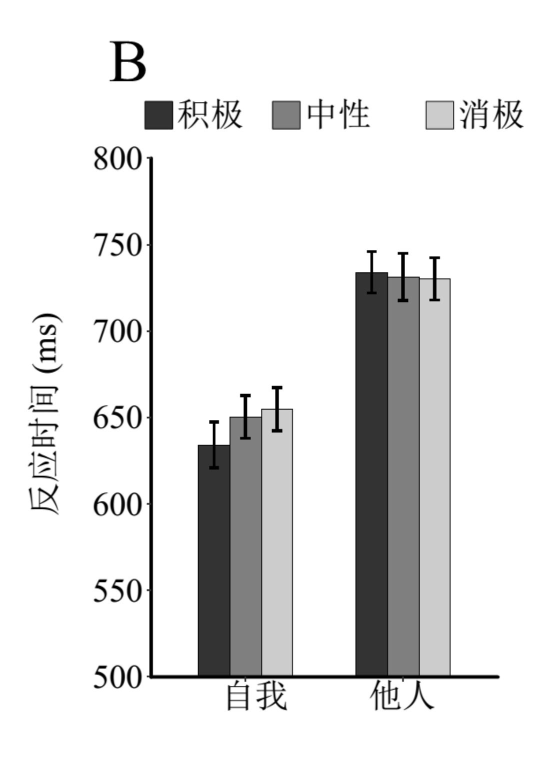


图 3.3　研究 5 中不同条件下的 *d*'（A）及反应时间（B）的均值，误差线为正负一个标准误。

3.3.4　讨论

研究 5 中，使用自我与他人作为联结的学习的任务，而将道德信息作为任务无关的任务，但仍然发现了积极道德信息对自我联结匹配任务的促进作用。这一结果为道德自我的内隐参照假设提供了支持，表明仅当积极的道德信息与自我同时出现时，即可促进知觉的加工。这个结果进一步表明，在知觉加工中，可能存在道德自我的内隐参照加工。

3.4　研究 6：自我相关性的内隐调节作用

3.4.1　引言

如前所述，研究 4 的结果表明，道德信息在知觉加工中的优势作用，可能是由于道德自我的内隐参照加工。研究 5 的结果进一步表明，即便在内隐的条件下，道德效价也能够调节自我相关信息的知觉加工优势。而本实验的目的，是为了进一步探索，自我参照信息是否会调节到道德效价的知觉加工优势。

本研究以道德联结任务作为外显任务，以自我刺激作为任务无关刺激，检验自我信息的呈现对道德联结的调节作用。即被试将首先学习道德效价与几何图形之间的联结，而自我相关的信息将作为任务无关的刺激呈现在图形之中，以检验自我相关信息作为与任务无关的刺激时，是否会影响到被试在道德匹配任务上表现。我们预期，与研究 5 一致，当积极的道德信息与自我同时呈现时，会促进知觉加工。

3.4.2　方法

3.4.2.1　被试

34 名大学生参加了本次实验，其中女性 26 名，平均年龄为 21.26 ± 1.78。所有的被试均是右利手，视力或者校正视力正确。实验前，所有被试签署《知情同意书》。2 名被试的数据由于总体的正确率接近随机水平而被排除出分析。最终有效数据为 32 名被试（25 名女性），平均年龄为 21.28 ± 1.76。

3.4.2.2　被试设计

与研究 4 相同，本实验为 2× 3× 2 完全被试内实验设计。第一个变量为自我相关性，包括自我和陌生人两个水平；第二个自变量为道德效价，包括积极、消极和中性的道德水平；第三个自变量为匹配程度，包括匹配和不匹配。

3.4.2.3　实验材料

本实验的实验材料与研究 4 的参数相同，但是使用三个几何图形：正方形、圆形和梯形。人物标签是：好人、常人和坏人。

与研究 1 相同，被试首先学习三个人物标签与几何图形之间的对应关系。开

始练习前，告诉被试每个人物标签的意义，比如“好人”代表的是道德高尚的人；而“坏人”则代表的是道德败坏之人；“常人”代表的是道德水平处于平均水平的人。被试看到各个人物标签与几何图形的对应关系并记住，然后被试开始练习。同样，人物标签与几何图形之间的对应关系在被试间是平衡的。

练习时，每个图形的中间均可能会有一个词：自己或者是他人。要求被试忽略图形中的文字，仅观察图形与正文标签之间的匹配关系。

3.4.2.4　实验程序

研究 6 实验程序与研究 5 相同，共 6 个 block，每个 block 有 120 个试次。

3.4.3　结果

研究 6 结果的分析方法与研究 5 相同。小于 200ms 的正确试次被删除，大约占全部试次的 6%。

对于 *d'* 的重复测量方差分析表明，自我参照与道德效价的交互作用边缘显著：$F(2, 62) = 3.123, p = 0.051$, $\eta_g^2 = 0.0049$。自我参照的主效应不显著：$F(1, 31) = 0.047$, $p = 0.8298$, $\eta_g^2 < 0.0001$；道德效价的主效应也不显著 $F(2, 62) = 1.925$, $p = 0.1545$, $\eta_g^2 = 0.0131$。

我们进一步进行简单效应分析，分别对自我和他人条件下进行重复测量的方差分析。在自我条件下，道德效价有显著的主效应：, $F(2, 62) = 3.674$, $p = 0.0311$, $\eta_g^2 = 0.0316$。事后比较发现，好人条件(2.609 ± 0.864)下的 d’大于坏人(2.28 ± 0.778), $t(31) = 2.268$, $p = 0.03048$, *Cohen's* $d_z = 0.4008$, 95% CI [-0.051 0.846]，也大于常人(2.343 ± 0.761)，$t(31) = 2.478$, $p = 0.01888$, *Cohen's* $d_z = 0.438$, 95% CI [0.036 0.835]；但是坏人条件和常人条件下则没有显著差异，$t(31) = -0.473$, $p = 0.63948$, *Cohen's* $d_z = -0.0836$, 95% CI [-0.463 0.287]。在他人参照的条件下，道德的主效应不显著：　$F(2, 62) = 0.638$，$p = 0.5317$，$\eta_g^2 = 0.0051$。

我们也检验了不同的道德效价条件下，自我出现和他人出现时的差异。结果表示，在坏人联结的条件下，出现自我时(2.28 ± 0.778)的 d’要低于出现他人条件 2.434 ± 0.869)): $t(31) = -2.393$, $p = 0.02294$, *Cohen's* $d_z = -0.4231$, 95% CI [-0.736 -0.08]; 但在常人条件下，出现自我(2.343 ± 0.761)和出现他人 (2.34 ± 0.844) 没有差异，$t(31) = 0.028$, $p = 0.97803$, *Cohen's* $d_z = 0.0049$, 95% CI [-0.351 0.371]；在好人条件下，出现自我(2.609 ± 0.864)和出现他人(2.484 ± 0.817)也没有差异，$t(31) = 1.582$, $p = 0.12373$, *Cohen's* $d_z = 0.2797$, 95% CI [-0.138 0.664]。

对匹配条件下的反应时进行重复测量方差分析发现，道德效价与自我参照之间的交互作用显著，$F(2, 62) = 5.108$, $p = 0.0088$, $\eta_g^2 = 0.0055$。道德效价的主效应显著，$F(2, 62) = 8.75$, $p = 0.0004$, $\eta_g^2 = 0.0728$；$F(1, 27) = 103.946$, $p < 0.0001$, $\eta_g^2 = 0.2943$；自我参照的主效应不显著，$F(1, 31) < 0.001$, $p = 0.9949$, $\eta_g^2 < 0.0001$。对于不匹配的试次的重复测量方差分析表明，道德效价的主效应显著，$F(2, 62) = 6.914$, $p = 0.0019$, $\eta_g^2 = 0.0144$。自我参照效应不显著，$F(1, 31) = 0.749$, $p = 0.3936$, $\eta_g^2 = 0.0004$；交互作用也不显著，$F(2, 62) = 0.064$, $p = 0.9384$, $\eta_g^2 = 0.0001$。

我们进一步对出现自我和出现他人条件下的数据进行重复测量方差分析。对于出现自我的条件，我们发现了显著的道德效价的主效应：$F(2, 62) = 11.474$, $p = 0.0001$, $\eta_g^2 = 0.1028$。配对样本 t 检验表明，好人条件下的反应时间 (670 ± 64)快于坏人条件 (715 ± 58): $t(31) = -3.875$, $p = 0.00052$, *Cohen's* $d_z = -0.6849$, 95% CI [-1.154 -0.238]，也快于常人条件(709 ± 57): $t(31) = -4.524$, $p = 0.00008$, *Cohen's* $d_z = -0.7997$, 95% CI [-1.222 -0.350]。常人条件下的与坏人条件下的反应时间没有显著差异, $t(31) = 0.535$, $p = 0.59624$, *Cohen's* $d_z = 0.0946$, 95% CI [-0.259 0.454].

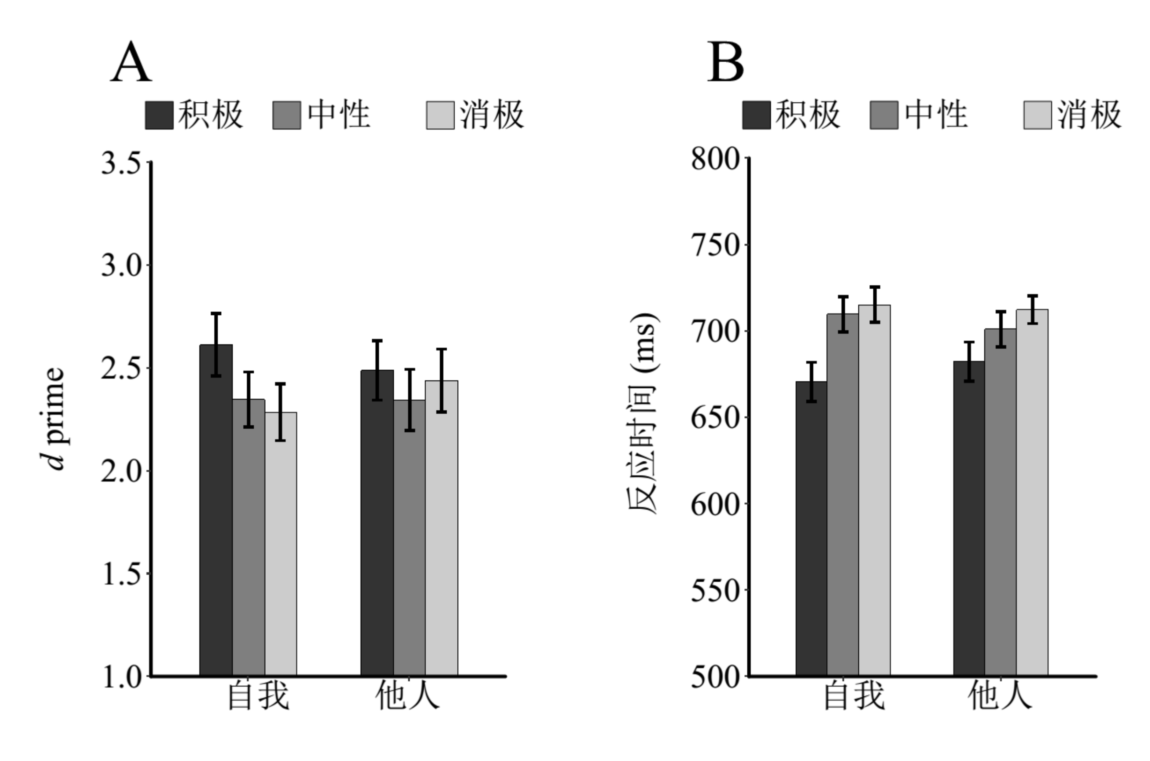


图 3.4 研究 6 中不同条件下的 *d'*（A）及反应时间（B）的均值，误差线为正负一个标准误。

出现他人的条件，道德效价的主效应也显著： $F(2, 62) = 4.928$, $p = 0.0103$, $\eta_g^2 = 0.0476$. 好人条件的反应时间 (682 ± 65) 快于坏人条件(712 ± 45), $t(31) =$ -

3.007, p = 0.0052, *Cohen's* d_z = -0.5315, 95% CI [-1.0023 -0.1402]；也快于常人条件 (701 ± 58): t(31) = -2.032, p = 0.0508, *Cohen's* d_z = -0.3592, 95% CI [-0.734 0.0218]。但是坏人与常人之间的反应时间没有显著差异 t(31) = 1.149, p = 0.25932, *Cohen's* d_z = 0.2031, 95% CI [-0.171 0.553]。

我们也检验了不同道德条件下出现自我和出现他人在反应时间上的差异，结果显示，在好人的条件下，出现自我显著快于出现他人，t(31) = -2.588, p = 0.01455, *Cohen's* d_z = -0.4576, 95% CI [-0.7707 -0.1506]；在常人条件下，出现自我显著长于出现他人, t(31) = 2.598, p = 0.01422, *Cohen's* d_z = 0.4593, 95% CI [0.0536 0.8748]；在坏人条件下，出现自我与出现他人没有显著差异，t(31) = 0.537, p = 0.59502, *Cohen's* d_z = 0.095, 95% CI [-0.2778 0.4413]。

3.4.4　讨论

研究 6 中，我们使用了与研究 1 相同的联结学习任务，但是将自我与他人作为无关刺激放在几何图形之中，结果发现，在自我出现的情况下，道德上积极的条件反应时更快，反应的准确率也更高。这个结果表明，即便自我信息在与任务无关时，也能够调节对道德效价的知觉加工优势，进一步验证了积极的道德效价与自我之间的内在联系。

研究 6 与研究 5 的结果一起，表明道德自我之间有着稳定的内隐的联系。当积极的道德信息与自我信息同时呈现时，个体的知觉加工会变得更好。这些结果与研究 4 的结果也保持一致，进一步表明，道德信息在知觉加工中的优势，可能是由于道德自我内隐参照加工。即人们以道德自我作为锚定，与道德自我越接近的信息，加工越快，与道德自我越远的信息，加工越慢。

但是，道德自我对知觉加工的调节效应到底有多大？在反应时间和 *d* prime 上是否具有相同的效应？要回答这些问题，需要对现在的 6 个实验进行量化的综合，因此我们进行了研究 7，来对道德自我在知觉加工中的调节效应进行准确地估计。

3.5　研究 7：对六个实验的元分析

3.5.1　引言

上述的六个实验表明，积极的道德效价能够对知觉起到促进作用，这种促进作用可能是由于道德自我的内隐参照加工引起的。然而，考虑到研究中可能存在的误差变异，使用实验内的元分析扩大样本量，提高研究的统计效力，同时能够对研究的效应进行更加准确的估计(Goh, Hall, & Rosenthal, 2016; Ueno, Fastrich, & Murayama, 2016)。因此，本研究将对以上 6 个实验进行元分析。6 个实验中总共包括 186 名被试，对于较小的效应量的检验也能够达到 80%以上的统计效力。

3.5.2　方法

3.5.2.1　软件

采用 metafor 工具包对 *d'*和 RT 的效应量的元分析(Viechtbauer, 2010)。该工具包采用了 Cooper, Hedges, and Valentine (2009)中对 Cohen's d 进行元分析。

3.5.3　结果

元分析结果显示，积极的道德自我信息比中性的道德自我信息具有显著的加工优势，在 *d*'的效应量为 Cohen's d = 0.293, 95% CI[0.091 0.495]，在反应时上的效应量为 Cohen's d = -0.542, 95% CI [-0.746 -0.338]。同样，积极的道德自我相比于消极的道德自我也具有显著的优势，在 *d*'上效应量为 Cohen's d =0.366, 95% CI [0.169 0.563]，在 RT 上的效应量为 Cohen's d = -0.6076, 95% CI [-0.804 -0.411]。但是中性与消极的道德自我信息之间并没有显著的差异。

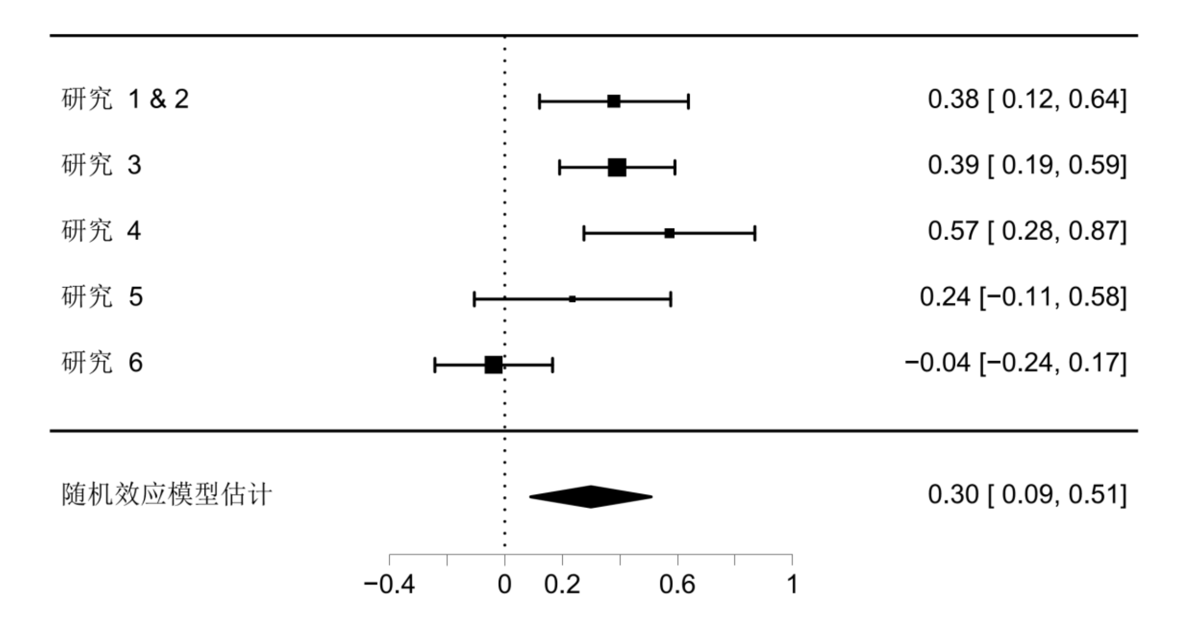


图 3.5　对于道德条件与不道德条件在 d' 上差异的元分析结果

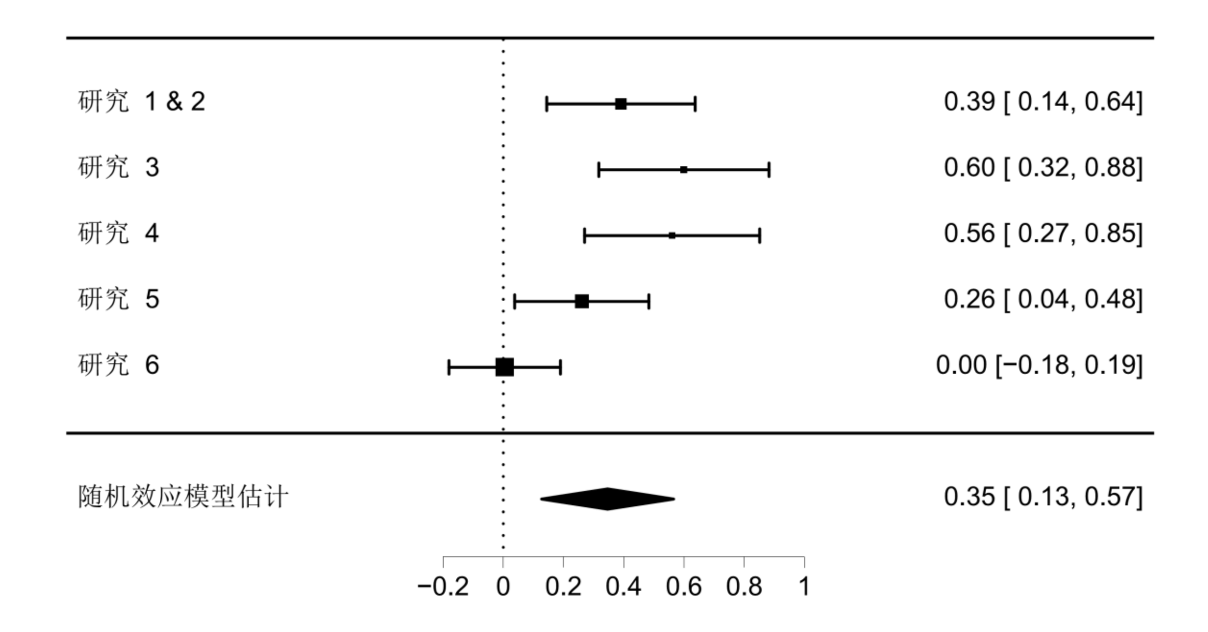


图 3.6　对于道德条件与中性条件在 d' 上差异的元分析结果

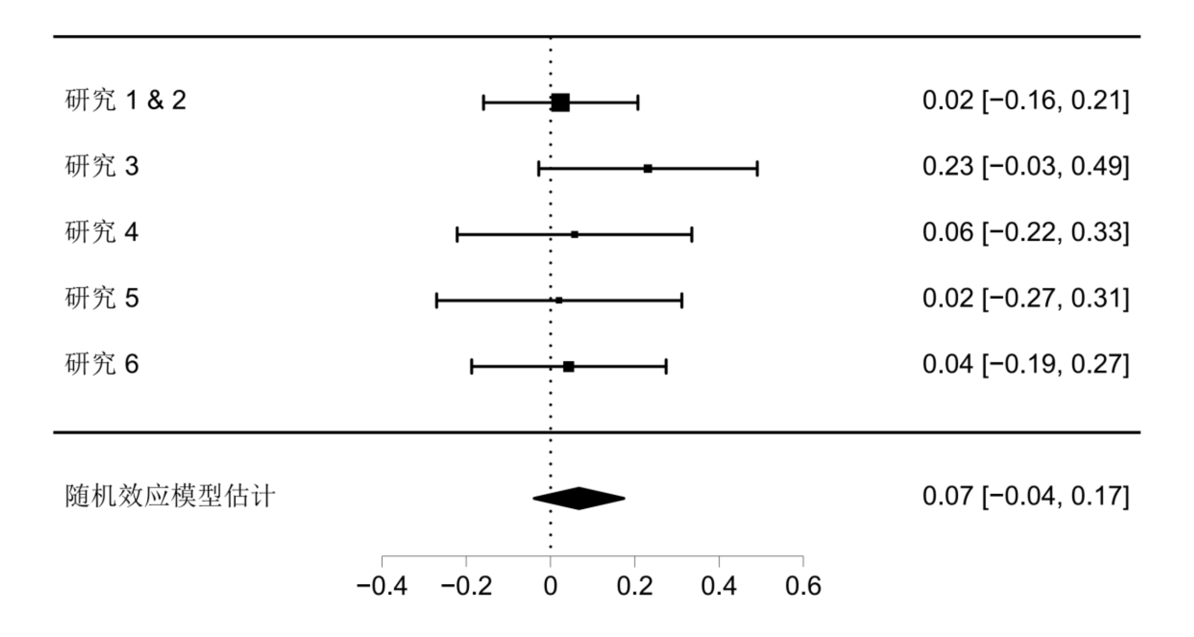


图 3.7　对于不道德条件与中性条件在 *d’* 上差异的元分析结果

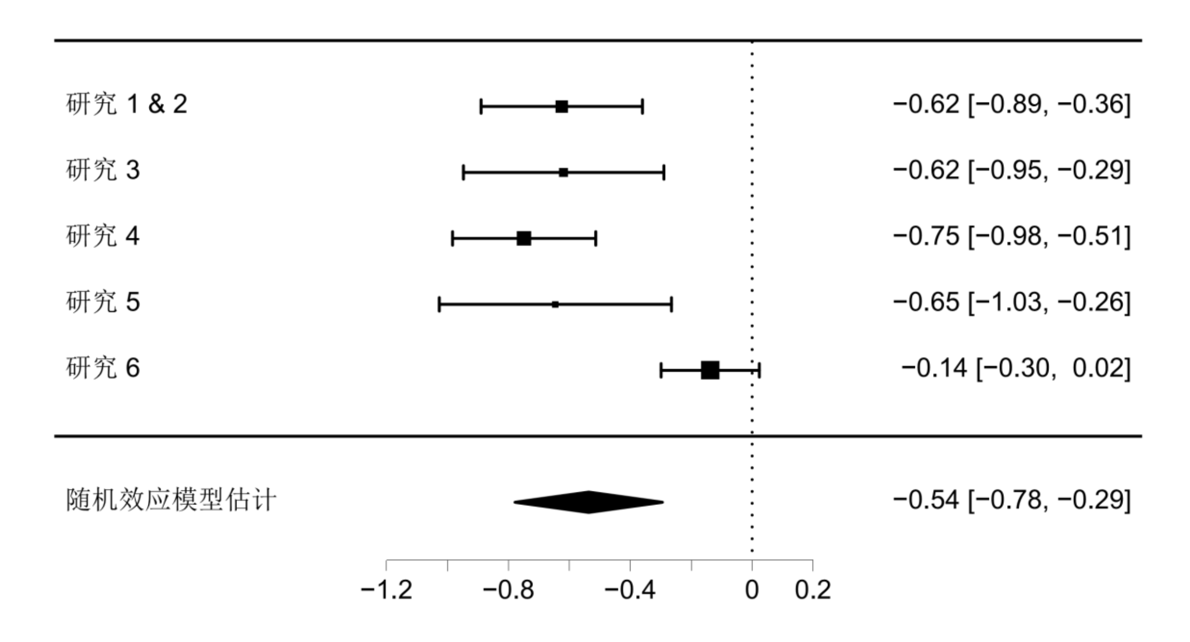


图 3.8　对于道德条件与不道德条件在 RT 上差异的元分析结果

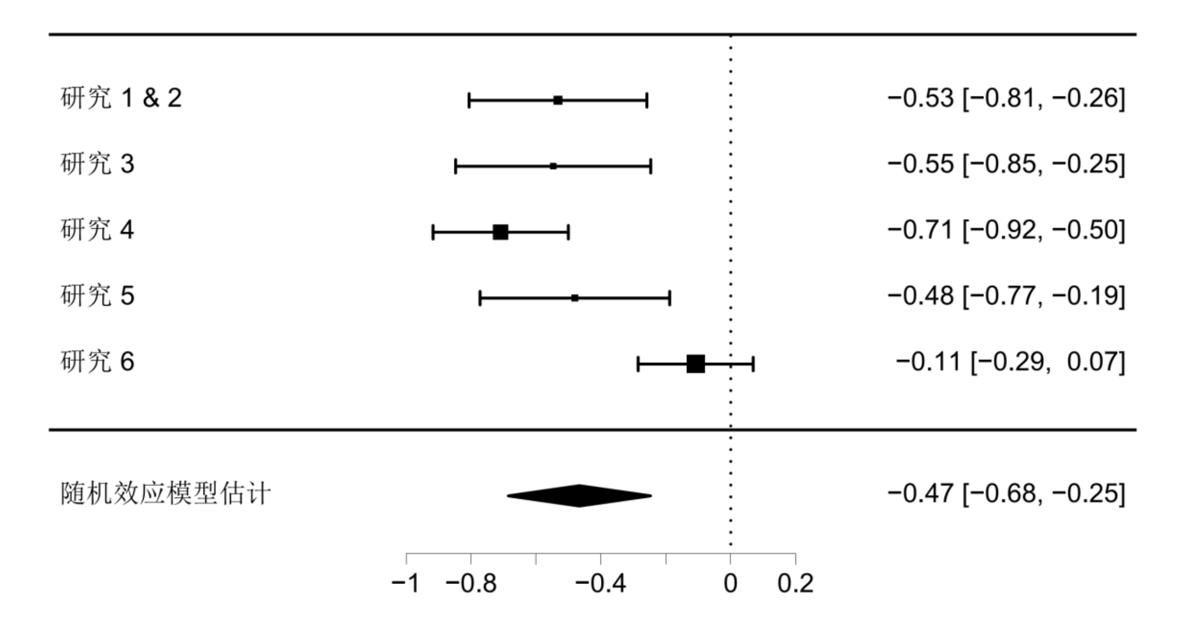


图 3.9　对于道德条件与中性条件在 RT 上差异的元分析结果

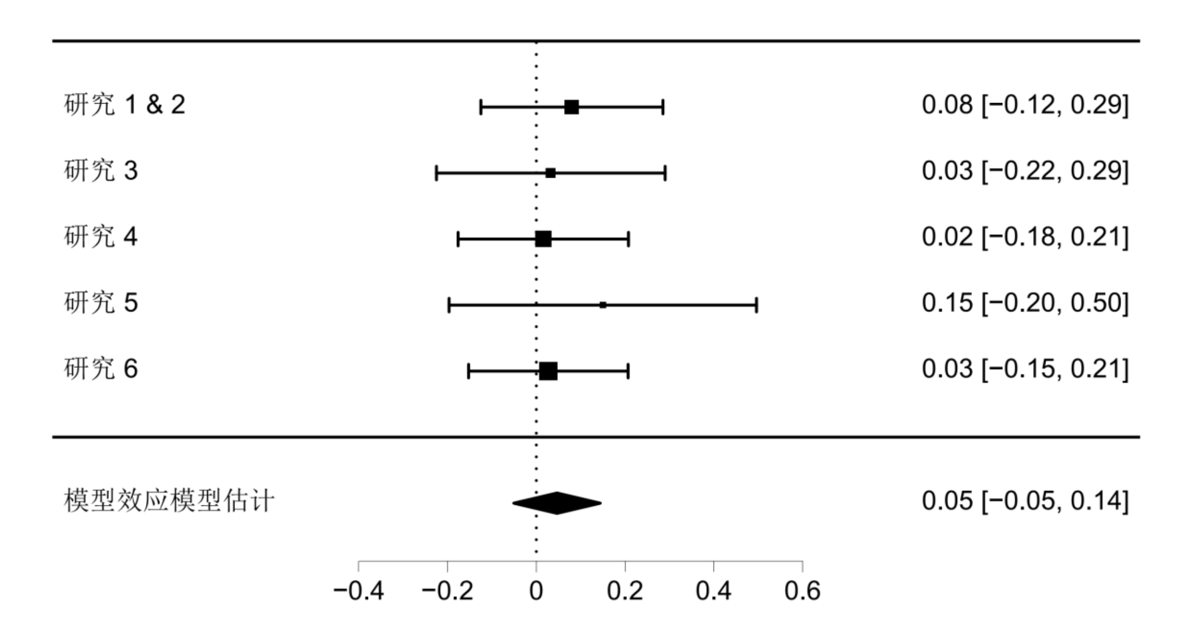


图 3.10　对于不道德条件与中性条件在 RT 上差异的元分析结果

3.5.4　讨论

对六个实验元分析的结果表明，道德自我对知觉起到调节作用，：积极的效价具有促进的，而消极的效应则与中性没有显著的差异。在 d'上，道德上积极刺激与中性和消极刺激之间的差异大约为 Cohen's $d = 0.3 \sim 0.35$，在反应时上，积极与消极和中性刺激之间的差异更大，约为 0.5。

这些结果表明，道德自我信息对知觉的调节作用稳定存在，而且其效应量也符合我们的常识，即具有中等的效应量。根据 Richard, Bond Jr, and Stokes-Zoota (2003) 对社会心理学一百多年的研究进行分析，社会心理学中的平均效应量约为 Cohen's $d = 0.4$，相关系数为 $r = 0.2$。本研究的结果表明，道德效价对知觉的调节作用属于正常的范围。

另一方面，由于正确率与反应时间上的效应量有所差异，反应时间能够更加敏感地探测出道德自我在知觉上的差异。从某种程度上讲，这也表明，将正确率与反应时间分开分析本身可能无法全面地考察被试的知觉过程。因此有必要使用模型分析的方法，将反应时与正确率结合起来，进一步细致地分析被试在知觉决策过程中的机制。

第四章　道德自我知觉优势的稳定性

4.1　本章引言

以上研究的结果表明，积极道德信息能够在知觉中起到促进使用，这种调节作用可能是由于道德自我的作用。但在对数据的分析上，我们采用了传统对反应时间数据的分析方法，即将每个被试在每种条件下的反应时间与正确率分开考虑，使用每种条件的平均反应时间作为其反应时间的指标，使用信号检测论的方式来分析其正确与错误反应的比例（*d'*）。在这种情况下，忽略了 RT 分布的形状以及错误反应的反应时间，也完全没有考虑反应速度和反应时间之间的关系(Wickelgren, 1977)。因此，使用传统的分析方法无法结合速度和准确度来综合考虑被试在实验中的反应情况。此外，标准的分析方法没有理论模型作为基础，这意味反应时间或者正确率上的差异无法与认知过程进行对应(Wagenmakers, Maas, & Grasman, 2007; 林少非, 2014)。

为了解决传统反应时分析的局限性，研究者提出了少模型来试图对行为进行分析。这些模型大多对反应时间的分布进行分析，从而获得比传统分析方法中只关注反应时间或者正确率的方法能够获得更加丰富的信息。在众多的模型之中，有两个模型得到的广泛运用：扩散决策模型（drift diffusion model, DDM）和指数高斯（ex-Gaussian）模型。

4.1.1　扩散模型（drift diffusion model，DDM）

对于有时间压力的二项迫选（2-choice forced task）任务，另外一个比较常用的模型是扩散模型（drift diffusion model）。

扩散模型将人们的简单反应时间过程当作是一个序列抽样的过程，即决策者不断地从刺激中抽取带有噪音的信息，信息逐渐累积，直到证据达到某个阈值，决策者做出决策(Forstmann, Ratcliff, & Wagenmakers, 2016)。这个模型考虑到决策过程中的速度-准确率权衡的矛盾，即一方面，如果有更多的信息，决策的质量会更高；另一方面，决策只有及时的时候他们才能被接受。事实上，在生物界内，这种速度-准确率的权衡广泛存在并且有着重要的影响，高海拔登山队员、棒球手、蜜蜂甚至单细胞生物如粘液菌都需要进行速度-准确率的权衡。

这种将决策过程看作是一个反复抽样过程的模型，叫做序列抽样模型(sequential sampling models)。基于序列抽样的思路，研究者们发展出了许多具体的

模型来对反应时间进行拟合，其中经典的序列抽样模型是累积子模型（accumulator models）(Van Zandt, Colonius, & Proctor, 2000; Vickers & Lee, 1998)。通常，序列抽样模型会假设决策者有一个预先存在的相对证据规则：一旦证据累积超过预定的阈值，即进行反应。在证据的累积是离散型数据的情境下，这类模型就是著名的随机游走模型（Random walk model）；在证据的累积是连续型数据的情况下，这类模型被称为扩散模型（drift diffusion model, DDM）。

随机游走模型最早可以追溯到概率论发展早期中的一个著名的问题——赌徒的毁灭（The gambler's ruin）。以下是赌徒的毁灭：

"假定有两个赌徒，A 和 B，他们将进行连续的赌博活动。假定他们的每次赌博相互之间是独立的、每个赌徒在开始时有一定金额的金钱（可能会不同），同时假定在每一次赌博中，A 赢的概率是固定。每次赌博之后，赢者可以得到对方一个单位的金钱。两人连续赌博，直到一人输光为止(Carazza, 1977; Feller, 1971)。"

这个情境带来的概率论问题是：如果已知两个赌徒开始时的金额数以及每次赌博中 A 赢钱的概率 p，那么，A 输光（"毁灭"）的概率是多少？预期从赌博开始到 A 输光，大约进行了多少轮赌博？(Feller, 1971)

在这个情境之中，金钱随着时间流动的过程可以使用随机游走过程来表征，且这个随机游走过程有两个边界：当达到其中一个边界时，表示其中一个赌徒破产。赌徒 A 在每次赌博中赢钱的概率参数 p 代表的是平均扩散速率（drift rate），如果 $p > 1/2$ 时，则随机游走的过程会偏向于赌徒 B 破产的边界。

如果赌博的赌注较小且 p 接近 1/2 时，这种随机游走过程也被称为布朗运动（Brownian motion）或者维纳扩散过程（Wiener diffusion process）。在数学领域，维纳扩散过程是一个随机微分方程的一个原型例子(Smith, 2000)，应用于金融、热流和流体动力学。

在心理学中，对随机游走模型应用的兴趣始于 Stone (1960)，随后是来自 Laming (1968)与 Link and Heath (1975)的贡献。但是早期的模型中均未考虑到错误反应时的相对速度。为弥补这一缺陷，研究者扩散了早期的模型得到了现在被研究者称为 2FAC 任务中标准模型的扩散模型。

扩散模型(Ratcliff, 1978; Ratcliff & McKoon, 2008; Voss, Nagler, & Lerche, 2013)假定二选的决策是基于对有噪音证据的累积，从一个开始点出发，终止于一个决策的阈值，这个阈值与特定的决策或者选择相联系。图 2.1 代表的是对扩散模型示意图。图形显示扩散的过程本身是有噪音的，导致了决策可能有错误以及 RT 的变化。

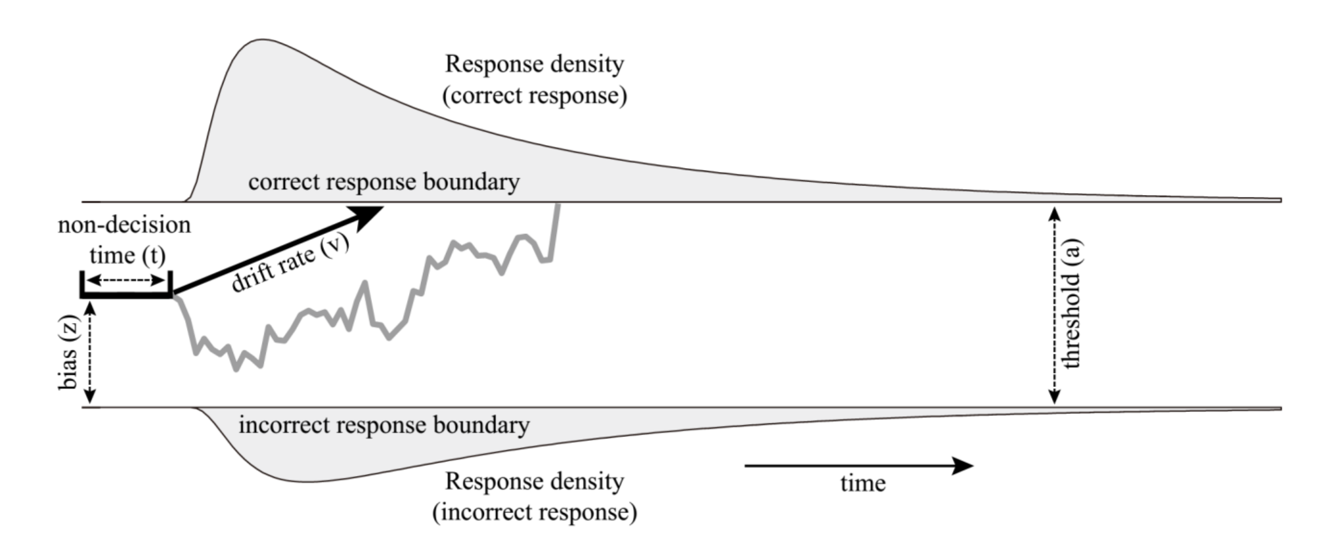


图 4.1 扩散模型示意图（修改自(Vandekerckhove, Tuerlinckx, & Lee, 2011; Wiecki, Sofer, & Frank, 2013)。

扩散模型有四个关键的参数。第一，扩散速率(drift rate) *v*，代表的是单位时间中证据累积的平均量，它是任务难度、被试能力、或者信息质量反映。第二，边界分离，也叫反应阈限（threshold）*a* 代表的是谨慎程度；边界分离的增加导致更少的错误（因为试次内的扩散噪音的影响减少）但是带来提是更慢的反应。因此边界分离表示的是一种速度-准确率的权衡。第三，起始偏见（initial bias）*z* 代表的对于两选一中某个选项的先验偏好。第四，非决策时间（non-decision time）*t* 代表的是决策之外的其他时间，如对刺激的编码和或者执行运作的时间。因此，总的反应时间是从起始点开始扩散到达边界的时间，加上非决策时间。

除了四个关键的参数外，扩散模型还考虑到了三种跨试次的变异：扩散速率、起始点和非决策时间。如果没有这三个试次间的变异，且边界与起始点的距离相同的话，则模型预测正确与错误的反应时间是相同的：在扩散速率高的情况下，正确与错误反应的反应时均变快，虽然正确率会变高；在低扩散速率的情况下，正确率下降，但正确反应与错误的反应时均减慢。因此，这种模型参数之下，快速反应的错误少、慢反应的错误多（相对于正确反应），导致错误反应比正确的反应慢（更多的解释见 Ratcliff (1978)和 Ratcliff and McKoon (2008)的图 4）。如果起始点有跨试次间的变异，模型预测错误的试次比正确的试次快(e.g., Ratcliff and Rouder (1998)中的图 2，以及 Wagenmakers, Maas, Dolan, and Grasman (2008)。当知觉加工过程的初始位置接近于正确反应的边界时，错误更少并且错误反应会更慢：因为加工过程需要更长的距离达到错误的边界。但是当知觉加工的过程在靠近错误边界开始时，则错误反应更多则且它们的反应时间很快，因为只需要很小的距离就到达错误

的边界。在已有的研究中，这两种模型均存在（有时会有交互作用，在简单的条件下更多快速的错误和在困难的条件下更多的慢错误，模型解释了为什么这些模型出现）。总体上，如果知觉加工的起始位置是混合的，则会有更快的错误反应。

通过维纳过程的数学公式，扩散模型可以根据正确与错误反应的分布，来计算概率密度。假定在时间点 x 时扩散模型终止于低的边界的概率记为$f(x|v,a,z)$，则该概率密度的分析解可以使用如下公式(Feller, 1968):

$$f(x|v,a,z)=\frac{\pi}{a^2}exp\left(-vaz-\frac{v^2x}{2}\right)\times\sum_{k=1}^{\infty}kexp\left(-\frac{k^2\pi^2x}{2a^2}\right)sin(k\pi z),(1)$$

上式中，x 表示当对低的边界做出反应的时间，v 是扩散速率，a 是边界分离的距离，z 是相对起始点（即 z 相对于 a 的起始位置，从 0 ~ 1 变化）。为了解决公式中的无限项，Navarro and Fuss (2009)提出了一个近似解，这个方法广泛地运用在扩散模型之中。

使用这些参数，扩散模型几乎为所有的快速反应时间任务中所观测到的数据提供了良好的解释。例如，扩散模型能够解释平均反应时间和选择可能性（错误和正确的反应）关系，包括它们如何随刺激难度而变化。此外，扩散模型可以解释反应时分布的解释为什么这些分布会随着实验条件的变化而变化。

虽然有研究者可能会质疑扩散模型的良好适用性来自其参数的灵活性，即模型对数据是过度拟合（overfitting）。但是，已有的研究表明，扩散模型的成功并非来自于过度的拟合(Wagenmakers, 2009)。首先，扩散模型的参数能够与实验设计中的条件对应起来，并且受到实验设计的调节。例如，只有扩散速率参数是与刺激的难度变化；只有边界分离是与不同的反应压力相对应的(Voss, Rothermund, & Voss, 2004)。使用实验设计来限制模型参数的变化，大大减少了模型的灵活性。第二，Ratcliff (2002)通过模拟分析表明，扩散模型无法解释看起来合理但是是人为的数据模式。第三，典型的实验设计可以“恢复”扩散模型的重要参数。最后，对模型参数特异性的测试表明，操纵特定的心理过程只会影响到相应的参数。例如，对任务难度的操纵对扩散速率影响最大，对时间压力的操纵对于边界分离的影响最大(例如，Voss et al. (2004)）。

扩散模型的出现，为分析时间压力之下的二选反应时间提供了非常好的模型。在认知神经科学得到发展之后，研究者发现扩散模型中的参数能够与神经数据非常好地结合起来，因此在认知神经科学领域也得到了广泛的应用(Forstmann et al., 2016; Ratcliff, Smith, Brown, & McKoon, 2016)。

为了推广扩散模型的使用，研究者们开发了多个工具来实现扩散模型。最近几

年来，已经有三个公开的工具包可以用来用来拟合完整的 DDM：基于 Matlab 的 DMAT (Vandekerckhove & Tuerlinckx, 2008)，独立的软件 fast-dm (Voss & Voss, 2008; Voss, Voss, & Lerche, 2015)和基于 Python 的 HDDM(Wiecki et al., 2013)。此外，还有简版的扩散模型 EZ-DDM 也可以实现(Wagenmakers et al., 2008; Wagenmakers et al., 2007)。此外，Wabersich and Vandekerckhove (2014)在实现 MCMC 的软件 JAGS 中将 DDM 分布作为一个部分。

4.1.2　指数高斯模型（ex-Gaussian model）

除了扩散模型之外，指数高斯模型的应用也比较广泛。指数高斯模型的提出，是为了弥补早期对反应时分布的缺陷。在早期对反应时间的分析中，主要是依赖于 Skewness 和 kurtosis，而没有任务数学模型(Sternberg, 1964, 1969)。但是，这种仅对反应时分布进行描述有两个主要的缺点：第一，抽样变异非常大，为了得到稳定的估计，需要几千个反应时间的数据；第二，容易受到极值的影响(Ratcliff, 1979)。因此即便收集了几千次反应时间，对这些参数的估计仍然可能被少量的极值反污染。

为了解决这两个缺陷，一种替代的方法是外显使用数学的模型来拟合反应时的分布。这样一方面避免了传统仅分布平均反应时间带来的对数据的错误表征，另一方面也提供了更多的信息来限制模型本身的。在这些尝试中，Hohle (1965)提出，反应时间的分布是正常分布与指数分布的卷积，即指数高斯分布（ex-Gaussian distribution）。对于大于 0 的 σ 和 τ，ex-Gaussian 分布的概率密度函数为：

$$f(t|\mu,\sigma,\tau) = \frac{1}{\tau(2\pi)^{1/2}} \exp\left[{\sigma^2}/{2\tau^2} - {(\mathrm{t}-\ \mu)}/{\tau}\right] \times \int_{-\infty}^{\left(\frac{\mathrm{t}-\mu}{\sigma}-\frac{\sigma}{\tau}\right)} exp\left(-\frac{y^2}{2}\right)^{dy} \quad (1)$$

其中 t 是时间，μ 是高斯分布的均值，σ^2 是正态分布的变异成分，τ 是指数分布的参数（即是均值也是标准差）。直觉上，μ 反应了众数（mode），τ 反映了 RT 分布的尾数的大小。

Hohle (1965)提出，指数高斯分布中的指数成分反映了被试的决策过程，而高斯成分则反映了残余的过程。这种将指数高斯分布与心理过程的对应关系却未被证明。但是指数高斯分布对反应时间的拟合效果却比较好。

研究者一方面优化算法来对反应时进行更好地拟合，另一方面探索指数高斯模型所代表的心理学意义。Heathcote, Popiel, and Mewhort (1991)使用 SIMPLEX 算法对拟合进行了探索，发现 SIMPLEX 算法能够较好地拟合 RT 数据。Van Zandt (2000)检验了几种拟合的方法后发现，连续最大似然法（Continuous Maximum Likelihood, CML）方法对数据的拟合得到最小的变异和最无偏的估计。随后，

Heathcote, Brown, and Mewhort (2002)提出使用分位数最大似然法（Quantile Maximum likelihood, QML），能够更好地拟合反应时分布。Cousineau, Brown, and Heathcote (2004)比较了几种最大似然法对 RT 的拟合之后发现，各种方法均能较好地对 RT 数据进行拟合。

除了指数高斯分布之外，还有一些类似的分布也被提出拟合反应时分布，例如 Weibull, Gamma, Wald, and ex-Wald。但是指数高斯分布具有模型的简洁性。如前所述，指数高斯函数的三个参数即可以反映整个分布的特点。虽然有一些争议，但是指数高斯分布的参数与心理过程也有一些相联系的证据。例如，高斯分布的均值的变化可能反映了初始刺激的加工，对刺激－反应联系的提取(Moutsopoulou & Waszak, 2012)或者在不同条件下噪音的处理(Balota, Yap, Cortese, & Watson, 2008)。相反，指数函数参数的变化可能反应了决策的过程(Hohle, 1965)、冲突解决（如 Stroop 任务中一致与不一致试次的区别）(Tse & Altarriba, 2012)、持续注意(Zhou & Krott, 2016)或者在分类任务中对于更加需要注意意资源的刺激需要更多的时间来对他们进行知觉上的分类(Balota & Spieler, 1999; Schmiedek, Oberauer, Wilhelm, Süß, & Wittmann, 2007; Staub, 2010)。更重要的是，先前的研究使用指数高斯函数来检验了认知加工中的特点。例如，先前对自我面孔的研究中发现，对自我面孔的加工在不同的任务和面孔背景之下，反应时间的分布不变，仅反应时间变化，但是对于熟悉他人或者不熟悉的他人面孔，则不仅时间发生变化，而且分布也会变化(Sui & Humphreys, 2013)。

在本研究中，对实验结果的 RT 分布进行分析，从而得到关于不同条件下反应时分布的特别，将有助于进一步理解在知觉决策过程中，道德自我的加工优势到底是如何发生的。

4.1.3　本章研究计划

如前所述，认知心理学中对反应时间的分析已经相当深入，通过对反应时间分布的模型分析，对知觉决策过程进一步分解，探讨知觉中的决策过程。因此，在本章的研究中，我们采用扩散模型和指数高斯模型对行为数据进行分析，深入探讨道德自我对知觉的促进作用发生的机制，从而进一步理解道德自我在知觉早期加工阶段的影响。

本章包括两部分：第一部分（研究 8），使用扩散模型对先前研究中的部分实验数据进行重新分析，对探索道德自我在知觉匹配任务中知觉决策过程的特点。第二部分（研究 9），我们进行一个新实验，结合知觉匹配任务与知觉分类任务，进一步探索道德自我在知觉决策过程中的跨任务的一致性。

4.2　研究 8：道德自我在知觉匹配中作用——DDM 分析

4.2.1　引言

研究 1 到研究 7 表明，道德效价能够调节知觉过程，而道德自我信息是最具有知觉加工优势的刺激。然而，道德自我信息是如何影响知觉过程的却仍然不清楚。

为了进一步理解道德自我信息在知觉决策过程中的作用，我们对研究 4、研究 4 和研究 6 的数据使用模型分析方法，进行了进一步分析。

4.2.2　方法

4.2.2.1　数据来源

本研究的数据来源于研究 4、5 和 6，其数据分别为数据 1、2 和 3。

数据 1 包括 24 名被试的数据，平均年龄为 21.79 ± 2.08。实验数据为 2× 3× 2 完全被试内实验设计。第一个变量为自我相关性，包括自我和陌生人两个水平；第二个自变量为道德效价，包括积极、消极和中性的道德水平；第三个自变量为匹配程度，包括匹配和不匹配。三个变量组合包括 12 种条件，每个条件下有 60 个正式的试次。

数据 2 包括 28 名被试（15 名男性），平均年龄为 19.46 ± 1.2。实验数据为 2× 3× 2 完全被试内实验设计。第一个变量为自我相关性，包括自我和陌生人两个水平；第二个自变量为道德效价，包括积极、消极和中性的道德水平；第三个自变量为匹配程度，包括匹配和不匹配。三个变量组合包括 12 种条件，每个条件下有 30 个正式的试次。与数据 1 不同的是，被试在学习阶段只学习两种联结：自我与他人（具体信息见研究 5）。

数据 3 包括 32 名被试（25 名女性），平均年龄为 21.28 ± 1.76。实验数据为 2× 3× 2 完全被试内实验设计。第一个变量为自我相关性，包括自我和陌生人两个水平；第二个自变量为道德效价，包括积极、消极和中性的道德水平；第三个自变量为匹配程度，包括匹配和不匹配。三个变量组合包括 12 种条件，每个条件下有 60 个正式的试次。与数据 1 不同的是，被试在学习阶段只学习几何图形与道德信息的联结（具体信息见研究 6）。

4.2.2.2　数据分析

由于研究 4 中被试的平均正确率较低，而指数高斯分布对反应时间的分布主要是基于正确反应时间，因此不适用使用指数高斯模型进行分析，为了保持一致

性，本研究中仅使用扩散模型进行分布。

扩散模型（DDM）分析

如前所述，扩散模型将两选任务模拟成为一个证据累积的过程，其中感觉信息进入到信息加工的过程，主体根据任务的要求和内部的因素如记忆和价值来决定，当前的信息提供是选项 A 或者选项 B 的证据(Ratcliff, 1978; Ratcliff et al., 2016)。

在本研究中，我们对 DDM 的参数估计采用了层级贝叶斯的方法，这种方法在被试个体差异的固定效应和随机效应之间很好地进行了权衡(Vandekerckhove et al., 2011; Wiecki et al., 2013)，对于个体和群体的参数估计都能达到更准确的估计，尤其是当试次数量比较少的时候(Ratcliff & Childers, 2015)。具体而言， 我们采用最近开发的软件包 HDDM (http://ski.clps.brown.edu/hddm_docs; Wiecki et al., 2013)进行模型拟合。使用贝叶斯对参数进行估计时，不仅可以量化参数，更重要的是可以对估计过程中的不确定性进行量化。在 HDDM 对参数的 估计中，使用 Markov chain Monte-Carlo (MCMC)抽样方法来估计参数的后验分布。在本次实验数据的拟合中，所有的参数均采用 5 个 Markov chain Monte-Carlo (MCMC)链进行 20000 次抽样（抛弃前面的 2000 次）从而达到了稳定的效应。对 5 个链的 Gelman-Rubin 聚合检验(Gelman & Rubin, 1992)表明 20000 次抽样能够达到良好的聚合效果。

对于每个被试的每个条件的 DDM 参数，均根据一个正态分布进行分布（或者截断的正态分布，根据参数的区间边界），以组均值为中心，组的方差为变异。选择的分布是根据 20 个最好的 DDM 拟合数据的研究而得到的，包括了一系列的决策任务(Matzke & Wagenmakers, 2009)。

为得到最好的数据拟合的效果，我们变化的自由估计的参数来拟合不同的模型。由于我们对扩散速率感兴趣，所以所有的模型中均包括了参数 v。此外，我们将边界分离 a 固定，因为在我们的实验中，指导语没有变化，因此边界分离应该是在任务中保持不变(McCoy & Theeuwes, 2016)。为了选择最优的模型，我们使用 Deviance Information Criterion (DIC)作为模型比较的标准。DIC 广泛运用于层级模型的比较之中(Spiegelhalter, Best, Carlin, & Van Der Linde, 2002)，越小的 DIC 意味着最高的似然性以及最小的自由度（模型的复杂性）。

此外，我们还使用的 Posterior Predictive Check (PPC)来进一步检验模型的拟合效果。PPC 使用估计出来的模型参数，使用 MCMC 随机模拟出新的数据集，即为预测数据集。将预测数据集与我们在实验中得到的数据进行比较，即可以得知模型对数据的拟合程度如何。我们使用 MSE（mean squared error）作为 PPC 的标准，MSE 值越小，表明模型拟合得越好。

得到最优的拟合模型之后，我们通过分析贝叶斯后验分布的方法来检验假设。

具体而言，如果某个参数的 95%后验密度大于 0，或者两个参数差的分布的 95%的密度大于 0，我们则认为其是统计上显著的(Cavanagh, Wiecki, Kochar, & Frank, 2014; Frank et al., 2015)。注意的是，所有的假设检验均是在组的平均后验分布上进行的。

4.2.3　结果

4.2.3.1　模型比较

我们进行了模型比较，选择出最佳拟合的模型。模型比较的结果显示，对于三组数据来说，最佳拟合的模型均是三个参数可以自由变化的模型（见表）。随后，我们从最佳拟合的模型中提取出每个条件下不同参数的后验估计分布并进行比较。

表 4.1　模型比较中的 DIC 值和 MSE 值。

	模型	可自由估计的参数	DIC	MSE
数据1	**1**	***v, t, z***	**-2218.7**	**0.014509**
	2	*v, t*	-1916.9	0.015407
	3	*v, z*	-1395.0	0.015901
	4	*v*	-823.3	0.016788
数据2	**1**	***v, t, z***	**-1220.9**	**0.011646**
	2	*v, t*	-1103.5	0.012781
	3	*v, z*	-898.7	0.013321
	4	*v*	-697.3	0.013715
数据3	**1**	***v, t, z***	**-3652.2**	**0.008514**
	2	*v, t*	-3347.0	0.008941
	3	*v, z*	-3004.8	0.009194
	4	*v*	-2409.8	0.009689

注: v = 扩散率，a = 决策阈限，z = 初始偏见，t = 非决策时间。字体加粗的模型是最佳拟合模型.

4.2.3.2 数据 1 结果

扩散速率：道德效价的优势仅对于自我条件有效。对于自我条件而言，好我的扩散速率高于坏我和常我：P_{Bayes}(好我 ＞ 坏我) = 1，P_{Bayes}(好我 ＞ 常我) = 0.999，但是坏我和常我条件没有差异：P_{Bayes}(坏我 ＞ 常我) = 0.565。但是对于他人条件，好人、坏人与常人之间没有差异：P_{Bayes}(好人 ＞ 坏人) = 0.457，P_{Bayes}(好人 ＞ 常人) = 0.498，P_{Bayes}(坏人 ＞ 常人) = 0.543。自我的优势效应也存在于道德条件：好我的扩散速率也高于好人：P_{Bayes}(好我 ＞ 好人) = 0.996，但是常我的扩散速率慢于常人：P_{Bayes}(常人 ＞ 常我) = 0.959；坏我的扩散速率慢于坏人：P_{Bayes}(坏人 ＞ 坏我) = 0.954。

起始偏见：好人条件下的起始位置大于坏人条件和常人条件： P_{Bayes}(好人 ＞ 常人) = 0.961, P_{Bayes}(好人 ＞ 坏人) = 0.932，坏人与常人间也未显示出差异：P_{Bayes}(好人 ＞ 常人) = 0.682。在自我条件下，好我、坏我与常我之间均没有差异：0.1899 ＜ Ps < 0.764。此外，好人的起始偏见也大于好我：P_{Bayes}(好人 ＞ 好我) = 0.915，但另外两种道德效价之下自我与他人无差异：P_{Bayes}(常我 ＞ 常人) = 0.698, P_{Bayes}(坏我 ＞ 坏人) = 0.262。

非决策时间：好我与好人的非反应时间适于其他条件（见图 4.2），其中，好我与常我之间接近显著：P_{Bayes}(好我 ＜ 常我) = 0.921，好人与坏人之间的差异显著：P_{Bayes}(好人 ＜ 坏人) = 0.958。而好人与坏人之间的差异不显著，P_{Bayes}(好我 ＞ 好人) = 0.561。表明其他条件下的非反应时间没有明显的趋势。

我们对不匹配条件下的数据也进行了同样的模型分析，采用与匹配试次相同的模型。结果表明，在扩散速率和非决策时间上，不匹配条件下并示显示出明显的模式（Ps < 0.863）。但是，在初始偏见（z）中，当图形为道德自我时，反应更加偏向于错误的反应（即按匹配键）：P_{Bayes}(好我 ＜ 坏我) = 0.959; P_{Bayes}(好我 ＜ 常我) = 0.975。但是坏我与常我没有差异，P_{Bayes}(坏我 ＜ 常我) = 0.423。在他人条件下，好人的初始偏见小于坏人人和常人，P_{Bayes}(好人 ＜ 常人) = 0.990, P_{Bayes}(好人 ＜ 坏人) = 0.939，坏人与常人条件也没有差异，P_{Bayes}(坏人 ＞ 常人) = 0.233。此外，三种道德条件下，自我与他人均没有差异：P_{Bayes}(好我 ＜ 好人) = 0.712，P_{Bayes}(常我 ＜ 常人) = 0.185，P_{Bayes}(坏我 ＜ 坏人) =0.354。表明对于道德的图形有一种“是”的反应倾向，对于不道德的图形，则有一种相反的趋势。

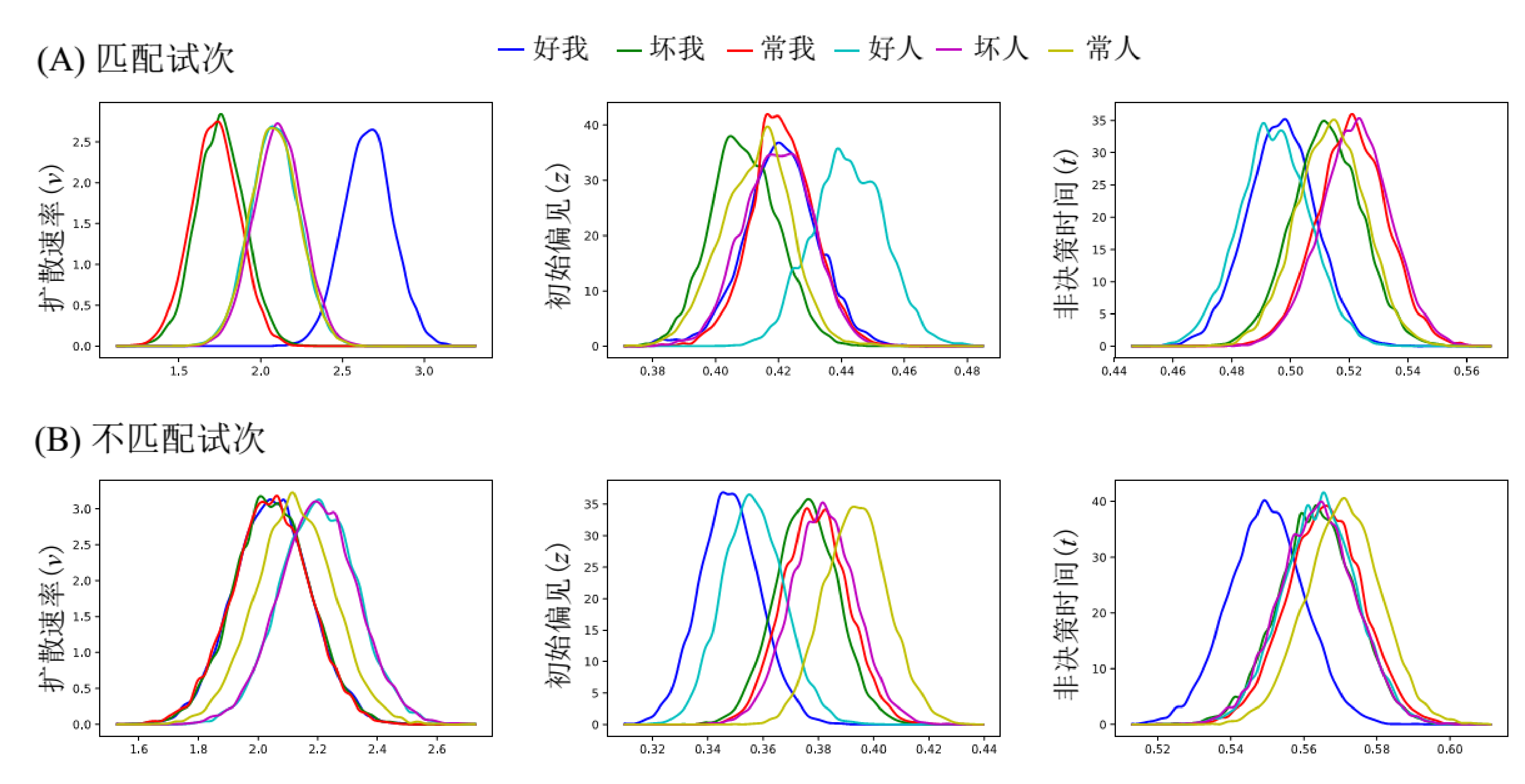


图 4.2　数据 1 中匹配试次（A）与不匹配试次（B）各条件下扩散速率（v），起始偏见（z）和非决策时间（t）的平均后验分布。

4.2.3.3 数据 2 结果

扩散速率：道德效价在于自我条件下，效应边缘显著：好我的扩散速率高于坏我：P_{Bayes}(好我 > 坏我) = 0.9172，好我与常我没有差异：P_{Bayes}(好我 > 常我) = 0.6758，坏我和常我条件没有差异：P_{Bayes}(坏我 > 常我) = 0.8189。但是对于他人条件，好人、坏人与常人之间没有差异：P_{Bayes}(好人 > 坏人) = 0.213，P_{Bayes}(好人 > 常人) = 0.6689，P_{Bayes}(坏人 > 常人) = 0.8941。自我参照的优势显著：好我的扩散速率也高于好人：P_{Bayes}(好我 > 好人) = 1，常我的扩散速率快于常人：P_{Bayes}(常我 > 常我) = 0.9981；坏我的扩散速率快于坏人：P_{Bayes}(坏我 > 坏人) = 1。

起始偏见：坏我的起始偏见大于坏人 P_{Bayes}(坏我 > 坏人) = 0.9767，其各条件均未发现差异 $Ps < 0.87$。

非决策时间：自我的各个条件均快于他人：P_{Bayes}(好我 < goodother) = 0.9804，P_{Bayes}(常我 < 常人) = 0.9602，P_{Bayes}(坏我 < 坏人) = 0.9961。在自我条件下，不同道德水平之间没有差异，$Ps < 0.651$;在他人条件下，不同道德水平之间也没有差异：$Ps < 0.8299$。

对不匹配条件下的数据也进行了同样的模型分析，采用与匹配试次相同的模型。结果表明，在扩散速率和非决策时间上，不匹配条件下并示显示出明显的模式。但是在起始偏见上，坏我大于常我 P_{Bayes}(坏我 < 常我) = 0.962；坏人大于常人和好人；P_{Bayes}(坏人 < 常人) = 0.988，P_{Bayes}(坏人 < 好人) = 0.9747。但是坏我与好我没有差异 P_{Bayes}(好我 > bad 好我) = 0.2856；坏人与坏我没有差异 P_{Bayes}(坏我 > 坏人) = 0.3928。

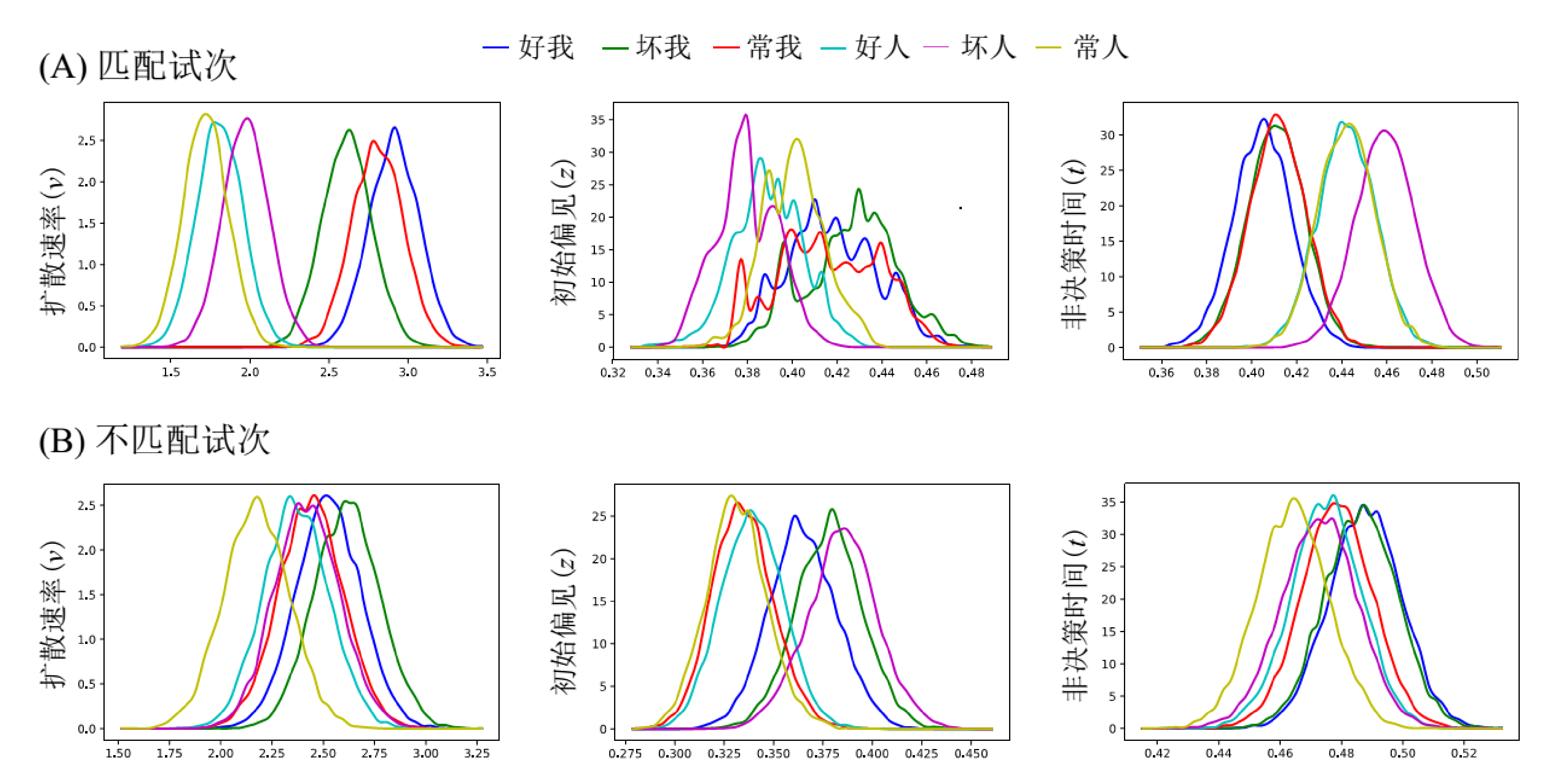


图 4.3　数据 2 中匹配试次（A）与不匹配试次（B）各条件下扩散速率（*v*），起始偏见（*z*）和非决策时间（*t*）的平均后验分布。

4.2.3.4 数据 3 结果

扩散速率：道德效价在于自我条件下，效应边缘显著：好我的扩散速率高于坏我：P_{Bayes}(好我 > 坏我) = 0.9374，好我快于常我：P_{Bayes}(好我 > 常我) = 0.9783，坏我和常我条件没有差异：P_{Bayes}(常我 > 坏我) = 0.2932。对于他人条件，好人有快于坏人和常人的趋势：P_{Bayes}(好人 > 坏人) = 0.9088，P_{Bayes}(好人 > 常人) = 0.8177，P_{Bayes}(常人 > 坏人) = 0.6734。各种条件下均没有表现出自我参照的优势：好我与好人，P_{Bayes}(好我 > 好人) = 0.781，常我与常人：P_{Bayes}(常我 > 常人) = 0.3608；坏我的扩散速率快于坏人：P_{Bayes}(坏我 > 坏人) = 0.728。

初始偏见：自我条件下，好我大于我坏我 P_{Bayes}(好我 > 常我) = 0.9505，常我大于坏我，P_{Bayes}(常我 > 坏我) = 0.9958，但是好我与常我没有差异 P_{Bayes}(好我 > 常我) = 0.263。在他人条件下，各种道德效价条件下没有差异，$Ps < 0.897$。此外，自我与他人在中性条件下有差异，常我的起始偏见大于常人 P_{Bayes}(常我 > 常人) = 0.9628，坏人的起始偏见有大于坏我的趋势，P_{Bayes}(坏我 < 坏人) = 0.9049。在道德条件下没有差异，P_{Bayes}(好我 > 好人) = 0.6346。

非决策时间：自我条件下，好我快于坏我，P_{Bayes}(好我 < 坏我) = 0.9813，也快于常我自我的各个条件均快于他人：P_{Bayes}(好我 < 常我) = 0.998。常我与坏我没有差异，P_{Bayes}(坏我 < 常我) = 0.795。他人条件下，则没有差异，$Ps < 0.6438$。此外，常我条件慢于常人条件，P_{Bayes}(常我 > 常人) = 0.9421。但好我条件与好人条件没有差异，好我的条件快于好人的条件，P_{Bayes}(好我 < 好人) = 0.8408；坏我条件与坏人条件也没有显著差异，P_{Bayes}(坏我 > 坏人) = 0.7778。

对不匹配数据的分析表明，在扩散速率、起始偏见和非决策时间上，不同条件之间均没有显著的差异（图 4.4）。

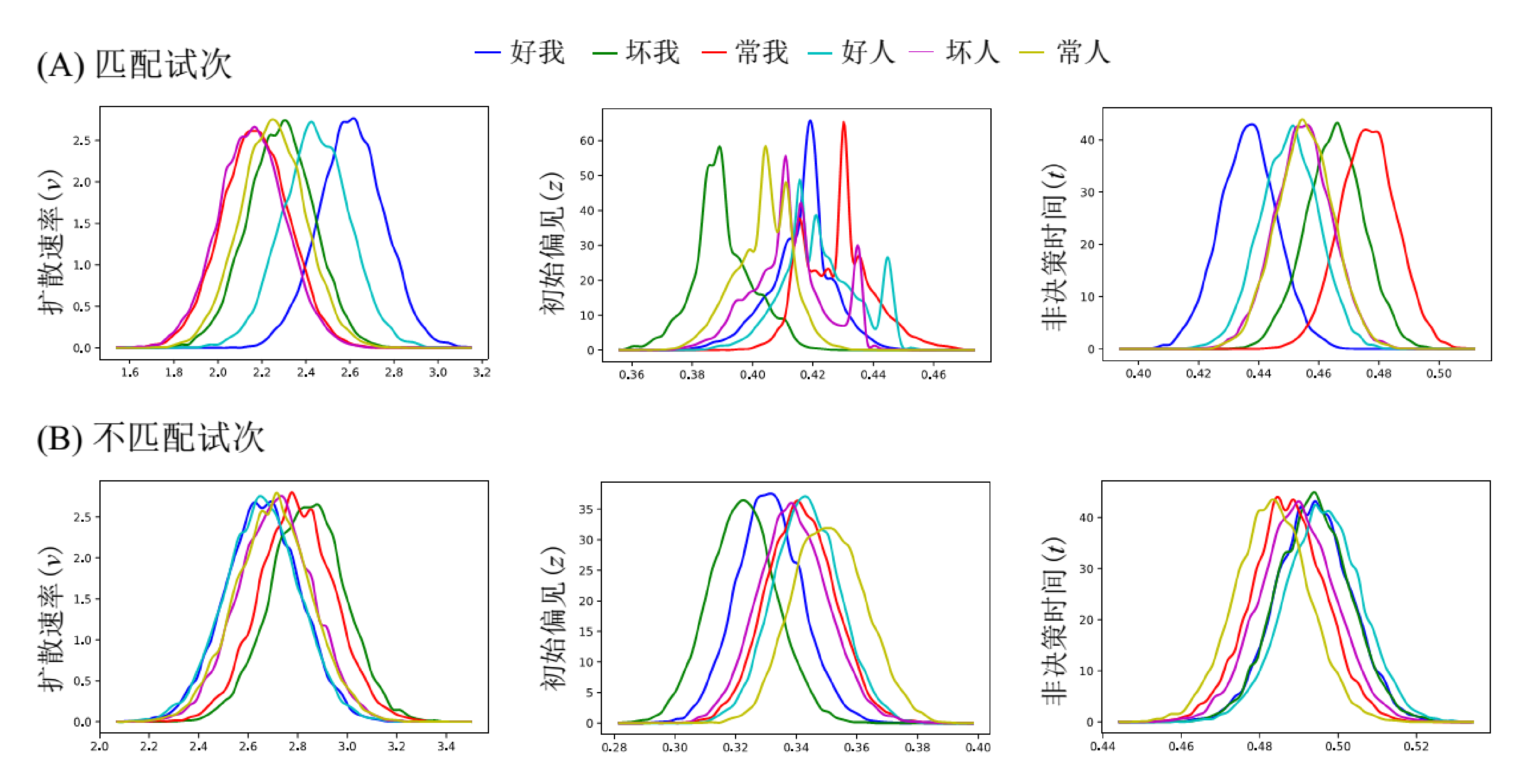


图 4.4　对数据 3 中匹配试次（A）与不匹配试次（B）各条件下扩散速率（v），起始偏见（z）和非决策时间（t）的平均后验分布。

4.2.4　讨论

通过对研究 4、5、6 的数据进行扩散模型的分析，我们发现，道德自我表现在各个任务中均表现出在信息累积率上的优势，这种优势非常稳定。而非决策时间上，则发现了道德自我相对于其他条件的优势则受到任务的调控。而在不匹配试次中，对于道德自我倾向于按匹配反应的倾向，也受到任务的调控。

先前知觉决策中的研究已经表明，扩散速率的变化，主要与刺激的信噪比呈正相关(Voss et al., 2004)。而在我们的研究中，在物理刺激保持一致的条件下，但由于刺激的社会信息的变化，带来了与先前物理刺激变化相似的结果，表明这种道德自我信息带来的知觉加工优势与物理刺激上的知觉加工优势有着类似的效应。

在非决策时间上，我们在外显任务时发现了道德自我与道德他人均有较短的非决策时间（数据 1 分析结果）。但是当自我联结变成外显任务而道德效价变成任务无关任务时，道德自我与自我的其他条件没有差异。相反，当道德效价任务变成外显任务但自我变成任务无关任务时，道德自我的非反应时间短于其他条件。这表明道德自我在非反应时间上的差异，可能主要来自于自我参照的效应。

初始偏见上，我们未发现稳定的效应，这可能与初始偏见参数的特性有关。先前的研究表明，初始偏见主要受到反应出现比例的影响，而在我们的研究中，匹配与不匹配出现的比例次数相同，因此可能被试并未对两种反应产生偏见。

总之，对研究 4、5、6 的数据进行扩散模型分析的结果表明，道德自我带来的优势，主要来自于扩散速率的加快，这种效应表明，道德自我这种社会信息的知觉加工优势，与物理刺激的知觉加工优势有类似的效应。此外，我们还发现这种知觉加工优势具有跨任务的稳定性。

但是，由于我们采用跨实验的分析，因此无法对道德自我的知觉加工优势进行因果推断。为了进一步检验这种跨任务的稳定性，我们进行了研究 9，直接检验道德自我信息在知觉加工中优势的跨任务稳定性。

4.3　研究 9：道德自我跨任务的稳定性

4.3.1　引言

如上所述，虽然研究 8 发现了道德自我的优势主要表现在扩散速度上，并且这种优势可能具有跨任务的稳定性。但是由于先前的研究未直接进行跨任务的比较，因此无法对道德自我的跨任务稳定进行直接推断。

为了进一步验证道德自我在知觉加工中优势的跨任务稳定性，在研究 9 中，增加了图形分类任务，探索道德自我在知觉匹配任务和知觉分类任务中是否有会同样的优势。

4.3.2　方法

4.3.2.1　被试

35 名大学生参加了本次实验，其中女性 14 名，所有的被试均是右利手，视力或者校正视力正确。实验前，所有被试签署《知情同意书》，实验结束后得到 60 元报酬。6 名被试的最终未被分析，其中 4 人因为程序错误而未完成实验，2 名被试的总体正确率接近随机水平而被排除出分析。最终有效数据为 29 名被试（13 名女性），平均年龄为 21.55 ± 1.99 年。

4.3.2.2　实验设计

本实验包括两个实验任务，其中知觉的匹配任务为 2× 2× 2 完全被试内实验设计。第一个变量为自我相关性，包括自我和陌生人两个水平；第二个自变量为道德效价，包括积极和消极两个道德水平；第三个自变量为匹配程度，包括匹配和不匹配。

第二个任务为知觉分类任务，该任务为 2× 2× 2 完全被试内实验。第一个变量为分类的标准，包括三个水平：按身份进行分类和按道德水平进行分类，第二个变量为图形的身份：自我或者陌生人；第三个变量为图形的道德水平：道德或者不道德。

4.3.2.3　实验材料与任务

本实验的实验材料包括四个几何图形：正方形、棱形、梯形和圆。人物标签是：好我、好人、坏我和坏人。

匹配任务　在任务开始之前，告诉被试从一系列陌生人名字中选择一下与自己性别相匹配的名字。然后告诉被试四个几何图形分别代表的是被试与这个陌生人的两个方面：道德的一面或者不道德的一面。例如，一个被试可能会告知："一个正方形代表的是你自己道德上高尚的一面，一个棱形代表的是你自己不道德的一面；梯形代表的是这个陌生人道德上高尚的一面，圆形代表的是陌生人不道德的一面。"四个图形与对应的标签相对应。图形与标签之间的顺序在被试间进行平衡。在这个阶段，不呈现图形与文字。

分类任务　在匹配任务之后，被试马上进行图形的分类任务。分类任务中 3.7 × 3.7 的图形呈现在屏幕中央，被试立刻进行一个图形的分类任务，要么根据图形的身份（自我 v. 陌生人），要么根据图形的道德属性（道德 v.不道德）。

所有的指导语和图形均是白色，呈现在灰色背景之上。图形的视角是 3.7×3.7 度，标签的神经是 3.6×1.6 度视角。图形与标签分别呈现在注视点（0.8×0.8 度视角）的上方和下方，注视点与图形和标签中央的距离为 3.5 度视角。

在这两个实验均在 PC 上，使用 22 寸 CRT 显示器（1024×768，100Hz）。刺激的呈现与反应的收集使用 Matlab 的 PsychToolbox-3 完成(Brainard, 1997; Pelli, 1997)。

4. 3. 2. 4　实验程序

被试在安静的测试室单个进行实验。实验分成两个阶段：被试首先学习图形与标签的对应关系，通过匹配任务，然后进行图形的分类任务。

在匹配任务中，每个试次首先以一个在屏幕中央的注视点 500ms 开始，然后图形－标签对呈现 100ms。一半的图形标签符合先前学习到的关系，属于匹配的试次，另一半则将图形与其他的标签进行配对，属于不匹配的试次。然后是空屏 1100ms。鼓励被试在这个期间进行反应，要求其反应又快又准。被试反应时需要使用右手的食指或者中指进行反应。空屏过去是 500ms 的反馈，如果被试反应正确，则给出正确，如果反应错误，则提示其反应太慢。每 block 结果之后，被试会被告知其整体的正确率。

在两个 block 的图形－标签匹配之后，被试开始图形分类任务。在这个任务中，被试的任务是根据预先设定的标准对图形进行分类判断。分类共有两个标准：身份标准：图形是属于自我还是属于他人；道德标准：图形是道德的还是不道德的？每个组块之后，在屏幕中央呈现指导语，告诉被试这个组块的分类标准。每个试次呈

现中央注视点 500ms，然后呈现图形 200ms。然后注视点 1000ms 作为反应的窗口。鼓励被试在这段时间内又快又准地反应。被试仍然需要通过右手的食指和中指进行按键反应。为了避免被试形成图形－反应键的联结，我们在被试间使用不同的反应键进行平衡。随后是 500ms 的反馈。试次间的间隔是 800~1200ms。

为了加强在图形分类任务中图形－标签的联结，在每个图形分类任务组块结束后，穿插一个图形标签匹配任务。因此对于匹配任务来说，被试首先完全 48 个试次的练习组块，仍然完成两个 120 试次匹配任务组块；随后在图形分类任务中完成 5 个穿插的组块，每个 48 试次。因此，每种条件下共 60 个实验试次。对于图形分类任务来说，被试完成 6 个 144 试次的组块，每个组块是三种分类标准之一（身份、道德效价和重要性）。因此，每个条件下有 72 个试次。

4.3.2.5　数据分析

由于分类任务中以重要性进行分类的数据在不同的被试间不相同，因此不纳入此次分析。

ANOVAs　本实验中，对反应时和正确率（或者 d'）进行重复测量的方差分析。在两个任务中，对 RT 的分析均排队了不正确的反应和快于 200ms 的反应（对于匹配任务，排队了 0.2%的试次，对于分类任务排队了 0.8%的试次）。我们报告了重复测量方差分析的效应量 generalized eta square (η_G^2) (Bakeman, 2005; Lakens, 2013)。此外，对于 *t* 检验的，我们使用 bootstrap 方法，报告了 95%的置信区间。这些分析均在 R 语言(R Core Team, 2016)中实现，其中方差分析使用了的是 EZ 工具包(Lawrence, 2015)，bootstrap 使用的是 BootES 工具包(Kirby & Gerlanc, 2013)。

指数高斯模型（ex-Gaussian）　对指数高斯模型的拟合采用 R 软件包中的“retimes”工具包进行指数高斯模型分析(Massidda, 2013)。具体而言，我们使用“timefit”函数来对每个被试每种条件下的正确反应时间进行拟合，得到每个被试在每种条件下的三个参数的值（μ、σ 和 τ），然后对这三个参数分别进行统计分析。“timefit”函数使用最大似然法来对数据进行拟合，其结果能够达到较好的效果(Cousineau et al., 2004; Van Zandt et al., 2000)。为了得到稳定的模型参数，“timefit”函数使用 bootstrap 的方法。我们对每个条件的估计均使用了 2000 次 bootstrap 的抽样。

扩散模型（DDM）分析　与实验 2a 相同，对三实验的结果也使用 HDDM 软件包进行扩散模型分析。由于本实验中包括两个实验任务，因此分别对匹配判断与分类判断中的数据进行扩散模型分析。

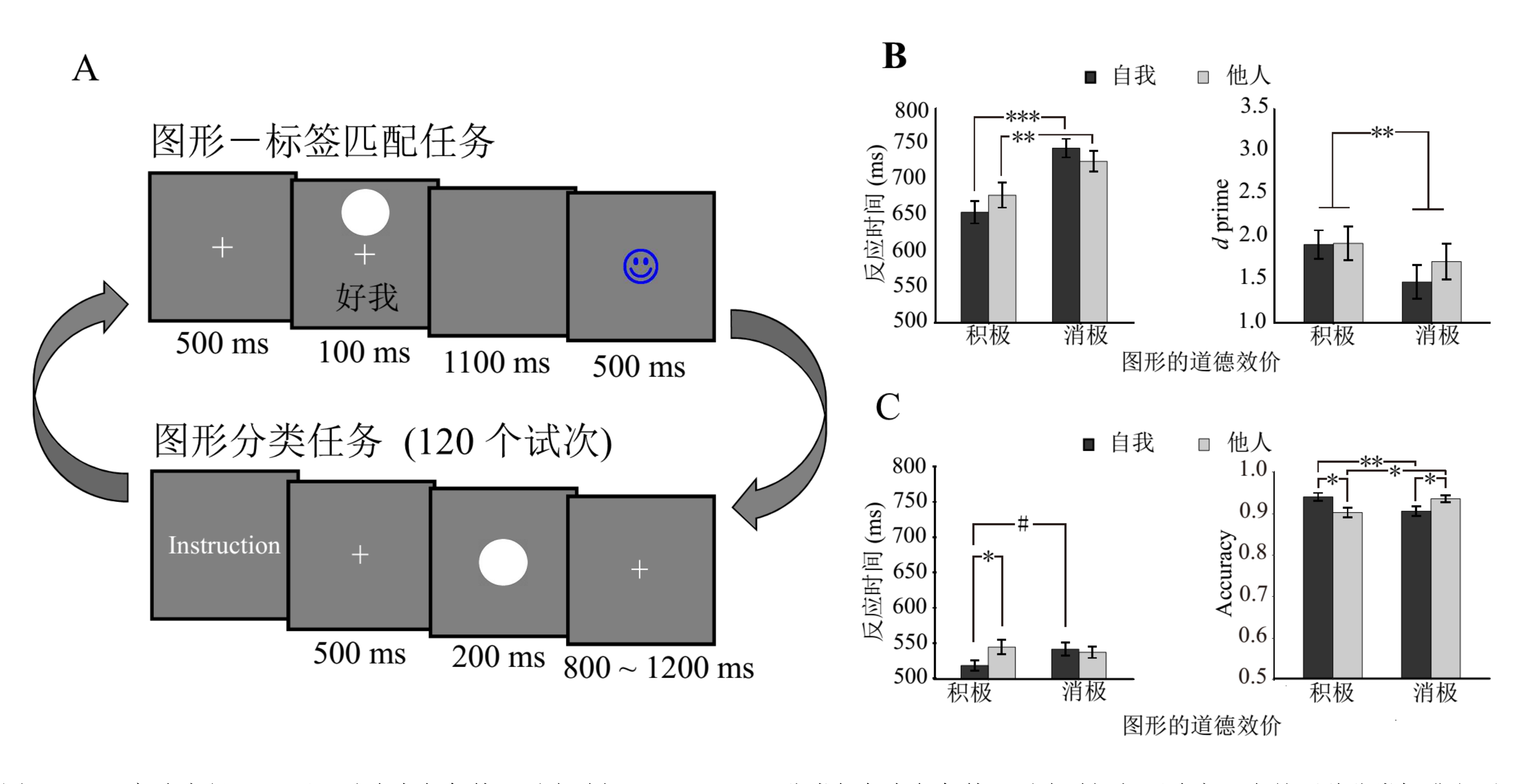


图 4.5 (A)实验流程；(B)匹配试次中各条件下反应时与 *d* prime ；(C)分类任务中各条件下反应时间与正确率（合并两种分类标准之后）。其中***表明在 0.001 水平显著，**表示在 0.01 水平显著，*表示在 0.05 水平显著，#表示边缘显著。

4.3.3 结果

4.3.3.1 知觉匹配任务

反应时间　对知觉匹配任务的反应时间进行重复测量方差分析表明，匹配的主效应显著：$F(1, 28) = 72.98, p < 0.0001, \eta_g^2 = 0.1082$。道德效价的主效应也显著 $F(1, 28) = 25.45, p < 0.0001, \eta_g^2 = 0.0439$。但是图形所属身份的主效应不显著：$F(1, 28) = 0.159, p = 0.693$。这些主效应伴随的是一个匹配、身份与道德效价的三项交互作用 $F(1, 28) = 9.42, p = 0.0047, \eta_g^2 = 0.007$。由于交互作用显著，　我们随后对匹配和不匹配的试次进行单独的分析。

对匹配试次的数据分析表明，道德效价的主效应显著，$F(1, 28) = 38.13, p < 0.001, \eta_g^2 = 0.153$，表明对积极效价更快的反应。身份的主效应不显著，$F(1, 28) = 0.078, p = 0.782$。在身份与道德效价之间有边缘显著的交互作用 $F(1, 28) = 3.939, p = 0.057, \eta_g^2 = 0.017$。随后的简单效应分析发现了对道德自我(655 ± 85 ms)比不道德自我(744 ± 70 ms)要更快：$t(28) = -5.67, p < 0.001$, $Cohen's\ d_z = -1.053$, 95% CI [-1.669 -0.3692]。同样，对道德他人(678 ± 93 ms)的反应也快于不道德他人(726 ± 77ms), $t(28) = -3.16, p = 0.004$, $Cohen's\ d_z = -0.588$, 95% CI [-1.01 -0.119]。尽管有一些交互的趋势，但是我们没有发现道德自我比道德他人显著更快 $t(28) = -1.536, p = 0.136$, $Cohen's\ d_z = -0.285$, 95% CI [-0.66 0.10]，也没有发现不道德自我比不道德他们显著更慢 $t(28) = 1.38$ $p = 0.1799$, $Cohen's\ d_z = 0.255$, 95% CI [-0.14 0.62]。

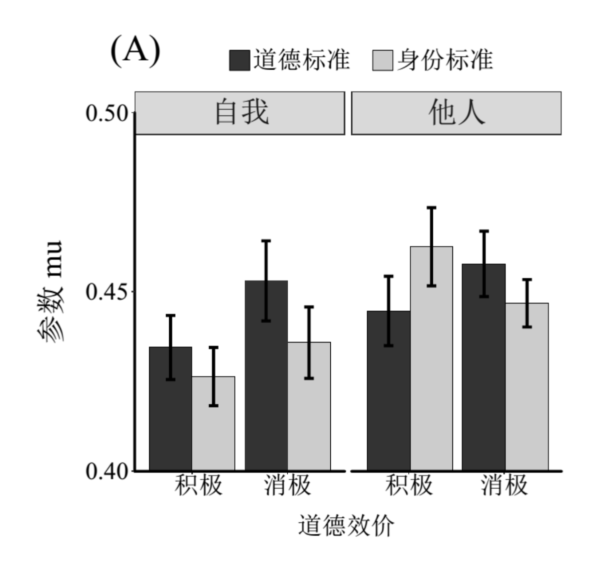
***d* prime**　对信号检测论的指标 *d* prime 的分析表明，道德效价的主效应显著，$F(1, 28) = 8.75, p = 0.006, \eta_g^2 = 0.025$，道德条件下的 *d* prime 要显著大于不道德条件。身份的主效应（$F(1, 28) = 1.12$，$p = 0.3$）以及道德效价与身份的主效应，$F(1, 28) = 1.12, p = 0.298$）均不显著。

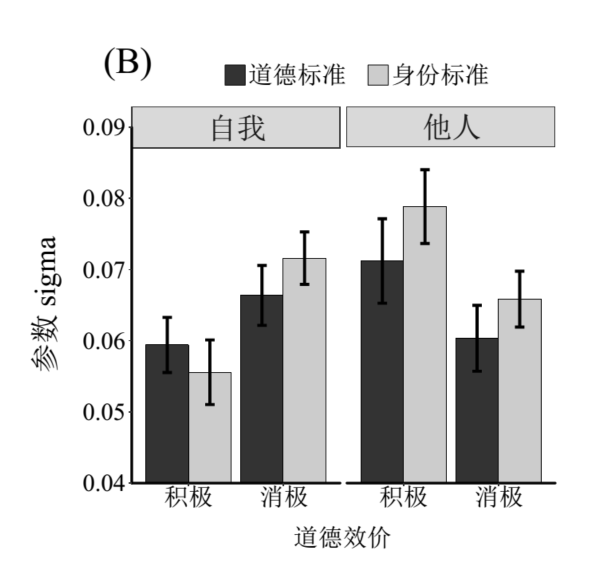
ex-Gaussian 分析　对指数高斯模型的参数估计进行了 2 × 2 × 2 的重复测量方差分析。对于参数 mu，我们发现了身份的主效应 $F(1,26) = 7.27, p = 0.012, \eta_g^2 = 0.015$，表明自我条件下的 mu (0.437 ± 0.0665) 显著小于他人条件 (0.452 ± 0.0628)。其他的主效应与交互作用均不显著：$Fs < 2.5, ps > 0.12$（见图 4.6A）。

对于参数 sigma，我们发现了身份的主效应 $F(1,26) = 4.67, p = 0.04, \eta_g^2 = 0.011$ 以及身份与道德效价的交互作用，$F(1,26) = 6.768, p = 0.015, \eta_g^2 = 0.038$。其他的主效应与交互作用均不显著，$Fs < 1.26, ps > 0.27$。为了进一步分析身份与道德效价的交互作用，我们将不同任务的数据合并，然后进行了 2 × 2 的重复测量方差分

析，结果表明身份的主效应显著 $F(1,26) = 4.67$, $p = 0.04$, $\eta_g^2 = 0.016$; 身份与道德效价之间的交互作用显著，$F(1,26) = 6.768$，$p = 0.015$，$\eta_g^2 = 0.056$。简单效应分析表明，对于自我条件，好我(0.058 ± 0.023)与坏我(0.069 ± 0.0226)之间的差异是边缘显著：$t(26) = -2.58$, $p_{adj} = 0.063$；对于他人条件，好人(0.0749 ± 0.0312)与坏人(0.0637 ± 0.024)之间的差异不显著 $t(26) = 1.96$, $p_{adj} = 0.24$。我们也比较了道德上积极与消极条件下，自我与他人的差异，结果显示，道德条件下，好我比好人有更小的 sigma，$t(26) = -2.94$, $p_{adj} = 0.027$；而对于不道德条件，坏人与坏我没有显著差异：$t(26) = 1.22$, $p_{adj} = 0.928$（见图 4.6B）。

对于参数 tau，我们没有发现任何显著的结果：$Fs < 1.51$, $ps > 0.229$ （见图 4.6C）。





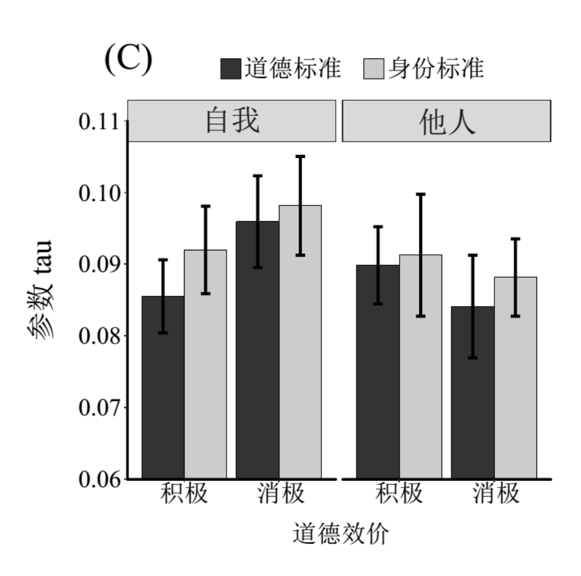


图 4.6 ex-Gaussian 参数随着分类任务、图形的身份以及图形的道德效价的变化。(A)参数 mu；(B)参数 sigma；(C)参数 tau。

HDDM 分析　　为了进一步检验上述效应在知觉决策中的哪个成分中出现，我们使用 DDM 来提取出知觉决策中不同的认知过程进行分析。首先，我们进行了一个模型比较的过程，选择出最佳拟合的模型，这些结果显示，最佳拟合的模型是三个参数均由数据来进行估计的模型（见表 4.2）

表 4.2 模型比较中的 DIC 值与 MSE 值。

模型	自由估计的参数	匹配任务		分类任务－道德标准		分类任务－身份标准	
		DIC	MSE	DIC	MSE	DIC	MSE
1	***v, t, z***	**-120.9**	**0.029657**	**-8052.6**	**0.003825**	**-6591.5**	**0.004325**
2	*v, t*	186.2	0.030882	-7700.2	0.004316	-6328.4	0.005063
3	*v, z*	475.8	0.032913	-7985.0	0.003796	-6418.6	0.004492
4	*v*	897.3	0.033218	-7539.9	0.004558	-6160.9	0.005251

注. . *v* = 扩散率, *a* = 决策阈限, z = 初始偏见, *t* = 非决策时间. 粗体的是最佳拟合模型.

随后，我们从最佳拟合的模型中提取出每个条件下不同参数的后验估计分布并进行比较。对于扩散速率来说，信息累积对身份与效价均敏感，其中道德条件下的扩散速率是最大的 P_{Bayes}(好我) > 坏我) = 1.0, P_{Bayes}(好人) > 坏人) = 0.9223. P_{Bayes}(好我) > 坏人) = 0.979。但是道德条件下的自我条件与他人条件的差异不显著：P_{Bayes}(好我) > 好人) = 0.710。此外，不道德自我条件是最慢的：P_{Bayes}(坏我 < 坏人) = 0.966。

对起始偏见（z）的分析表明，相对于他人条件下，对道德自我有一种反应的偏见：P_{Bayes}(好我 > 好人) = 0.920, P_{Bayes}(好我 > 坏人) = 0.940；但是在道德自我与不道德自我之间，没有这种差异：P_{Bayes}(好我 > 坏我) = 0.756。

对非决策时间（*t*）的分析表明，不道德自我的非反应时间要短于道德自我非决策时间：P_{Bayes}(坏我 < 好我) = 0.91。其他条件下的非反应时间没有明显的趋势（$P_{Bayes} < 0.850$）。

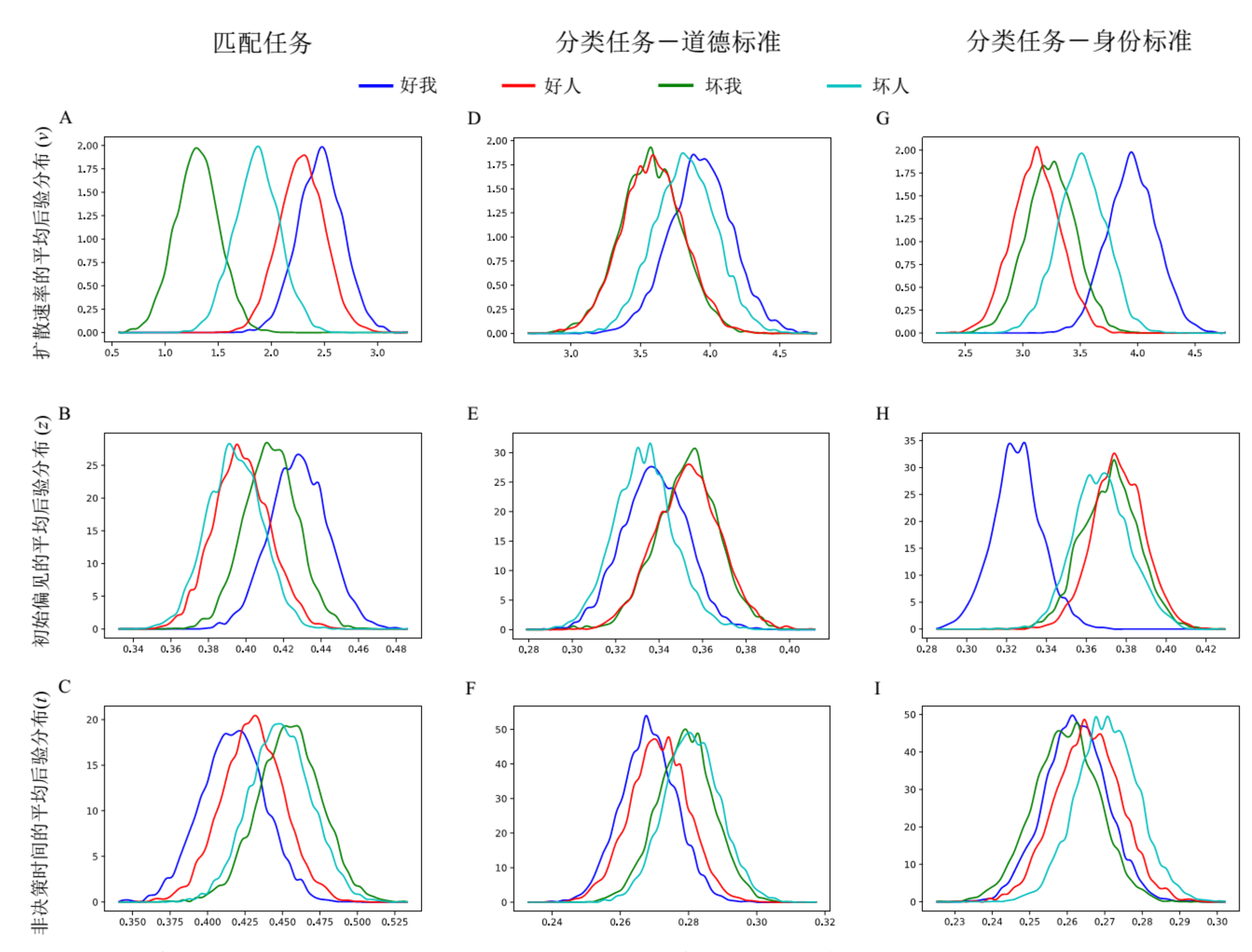


图 4.7 每个 DDM 参数在不同条件下的组平均后验分布。(A, B, C) 在匹配条件下的扩散速率 v、起始偏见 z、和非决策时间 *t*。(D, E, F) 道德效价分类条件下的扩散速率 v、起始偏见 z、和非决策时间 t。 (G, H, I) 身份分类任务条件下的扩散速率 v、起始偏见 z、和非决策时间 t.

我们对不匹配条件下的数据也进行了同样的模型分析，采用与匹配试次相同的模型。结果表明，在扩散速率和非决策时间上，不匹配条件下并示显示出明显的模式。但是，在初始偏见（z）中，当图形为道德自我时，反应更加偏向于错误的反应（即按匹配键）：P_{Bayes}(好我 < 坏我) = 1; P_{Bayes}(好我 < 好人) = 0.999, and P_{Bayes}(好我 < 坏人) = 0.96。表明对于道德自我的图形有一种"是"的反应倾向。反对，对于不道德自我的图形，则有一种相反的趋势：P_{Bayes}(坏我 < 坏人) = 0.99667; P_{Bayes}(坏我 < 好人) = 0.918667。

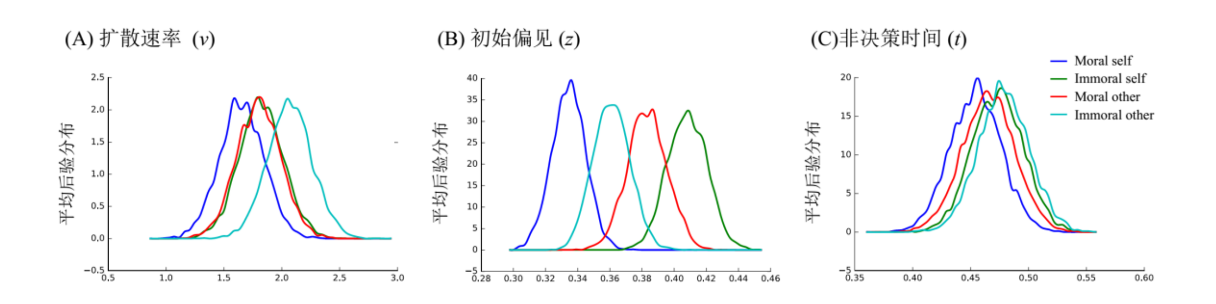


图 4.8 在非匹配试次下，模型参数在各个条件下的平均后验分布，(A)为扩散速率(v)，(B)为初始偏见(z)，(C)为非决策时间。

4.3.3.2 图形分类任务

反应时间　对反应时间进行 2×2×2 的重复测量方差分析表明，身份的主效应显著：$F(1, 28) = 8.09, p = 0.008, \eta_g^2 = 0.009$，表明对自我的反应快于对他人的反应。道德效价（$F(1, 28) = 3.553, p = 0.070$）和任务（$F(1, 28) = 0.12$，$p = 0.74$）的主效应均不显著。道德效价与任务类型的交互作用显著：$F(1, 28) = 7.57, p = 0.01, \eta_g^2 = 0.008$. 其他的交互作用均不显著：$Fs, ps > 0.120$ （见图 4.6）。

进行一步对不同身份条件的数据进行合并后，对不同分类任务下的数据进行单独分析。对于身份分类任务，道德效价的主效应不显著：$t(28) = 0.398$, $p = 1$, Cohen's d = 0.074, 95% CI[-0.322 0.427]. 按道德条件分类时，道德的主效应边缘显著, $t(28) = -2.65$, $p = 0.053$, Cohen's d = -0.492, 95% CI[-0.845 -0.065]，道德的图形分类快于不道德的图形。

正确率　对正确率数据进行同样的重复测量方差分析，表明身份与道德效价之间存在边缘显著的交互作用：$F(1, 28) = 4.043, p = 0.054, \eta_g^2 = 0.01$。结果显示，其他的主效应或者交互作用均不显著，$Fs < 2.89, ps > 0.1$（图 4.6）。

随后，我们对不同条件条件下的正确率数据合并之后，进一步检验道德效价与

身份的简单效应。对于自我图形的条件，在道德自我(0.923 ± 0.08)与不道德自我(0.884 ± 0.099)之间的差异显著：$t(28) = 3.594$, $p = 0.0049$, Cohen's d = 0.67, 95% CI[0.337 1.012]，但是在他人图形条件下，道德他人(0.88 ± 0.098)与不道德他人(0.895 ± 0.167)条件之间的差异不显著：$t(28) = -0.49$, $p = 1$, Cohen's d = -0.09, 95% CI[-0.75 0.312]。我们也检验了不同道德效价条件下自我与他人的正确率上的差异。对于道德上好的条件，在自我与他人之间的差异显著 $t(28) = 3.45$, $p = 0.0073$, Cohen's d = 0.639, 95% CI[0.126 1.091]。但是对于道德上坏的条件，则没有这种差异：$t(28) = -0.53$, $p = 1$, Cohen's d = -0.098, 95% CI[-0.641 0.318]。

HDDM 分析　与匹配任务的分析相似，我们对于分类任务的数据也使用 HDDM 进行拟合。模型比较的结果表明，自由估计扩散速率、初始偏见和非决策时间三个参数得到的最佳拟合的模型（见表 4.2）。然后，我们继续分析了最佳拟合模型的三个参数结果。

对于扩散速率而言，在身份分类任务中，道德任务大于其他条件：(P_{Bayes}(好我 > 坏我) = 0.987), (P_{Bayes}(好我 > 好人) = 0.993)，(P_{Bayes}(好我 > 好人) = 0.951)。表明在身份分类任务中，道德自我具有最快的信息累积速率。在基于道德效价的分类任务中，也有同样的结果模式，虽然证据没有这么强：P_{Bayes}(goo self > 坏我) = 0.896, P_{Bayes}(好我 > 好人) = 0.888, and P_{Bayes}(好我 > 坏人) = 0.833 （见图 4.7）。

与扩散速率的结果相似，在身份分类任务中，对道德自我有较强的起始偏见：P_{Bayes}(好我 < 坏我) = 0.991; P_{Bayes}(好我 < 好人) = 0.973; P_{Bayes}(好我 < 坏人) = 0.988；但是其他条件之间却没有显著区别 $P_{Bayes} < 0.704$。相反，在基于道德效价的分类任务中，没有发现图形的身份或者道德效价会引起初始的偏见：$P_{Bayes} < 0.812$（见图 4.7）。

对非决策时间的分析表明，图形的身份与效价在两种分类任务中均不影响非决策时间：$P_{Bayes} < 0.817$（见图 4.7）。

4.3.3.3 跨任务分析

为了进一步检验在匹配任务与分类任务上效应的稳定性，我们对两任务下的反应时、正确率、扩散速率和起始偏见进行了相关分析（见图 4.9）。具体而言，我们分类了两种任务下的自我偏见效应（即对自我相关信息的优势加工效应）和积极效价效应（即对积极信息的优势加工效应）。其中，自我偏见效应是通过道德自我与道德他人之间的差异；积极效价效应是通过道德自我与不道德自我条件之间的差异。

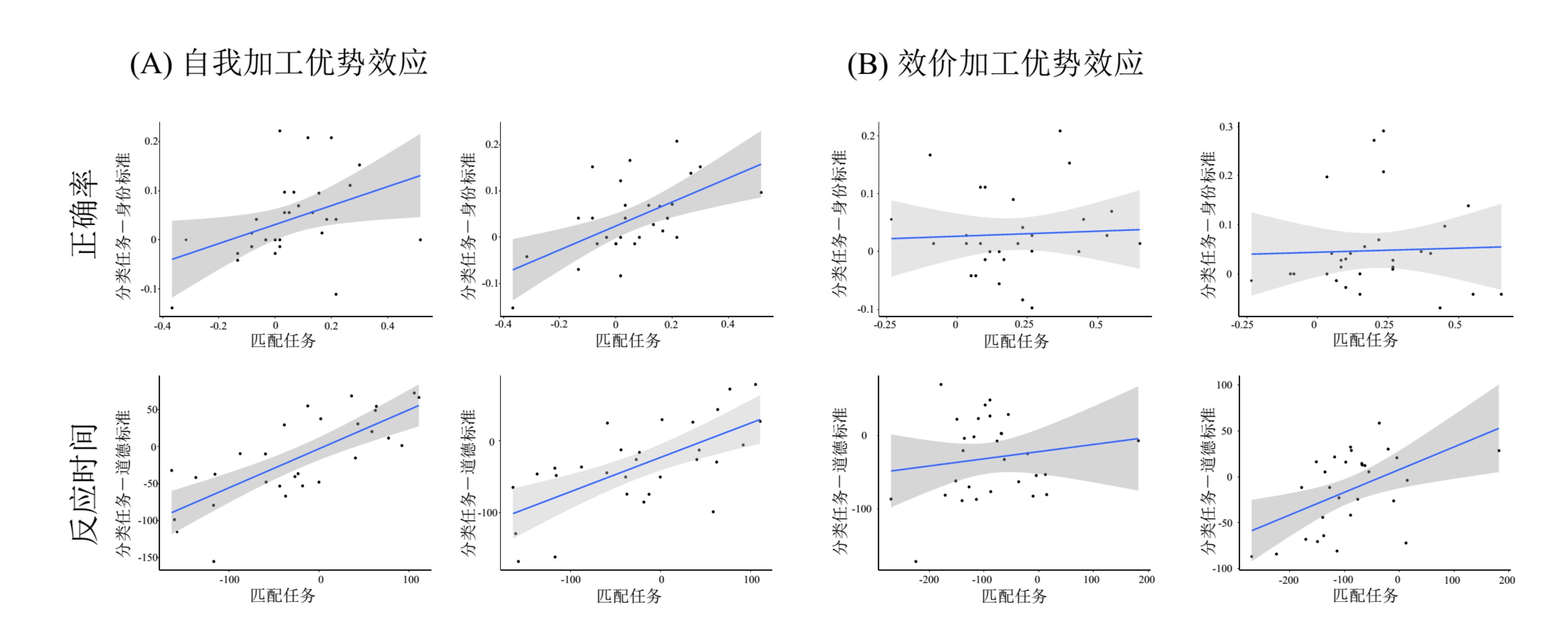


图 4.9　跨任务的相关分析。(A) 匹配任务与道德标准的分类任务中自我偏见的相关；(B) 匹配任务与道德效价分类任务中道德偏见的相关。

结果表明，匹配任务与道德标准的分类任务中自我偏见呈正相关，在正确率上的相关系数为 r = 0.4058, 95% CI [0.046 0.672], t(27) = 2.31, p = 0.028；在反应时间上的相关系数为 r = 0.756, 95% CI [0.539 0.879], t(27) = 6, p < 0.0001。同样，在匹配任务与身份标准的分类任务中自我偏见的也呈正相关，在正确率上的相关系数为 r = 0.575, 95% CI [0.264 0.777], t(27) = 3.65, p = 0.001；在反应时间上的相关系数为 r = 0.662, 95% CI [0.389 0.827], t(27) = 4.58, p < 0.0001（见图 4.9A）。但是我们在道德效价偏见上没有发现稳定的正相关（见图 4.9B）。

对扩散速率的相关分类也得到了类似的结果。在自我偏见中，匹配任务与身份分类任务的自我偏见的扩散速率呈正相关：r = 0.538, p = 0.003, 95% CI [0.214 0.756]；匹配任务与效价分类任务中自我偏见的扩散速率也呈正相关：r = 0.501, p = 0.006, 95% CI [0.165 0.733]。对于道德效价偏见而言，则没有显著的相关：匹配任务与身份分类任务：r = 0.219, p = 0.253, 95% CI [-0.16 0.542]；匹配任务与效价分类任务：r = -0.132, p = 0.493, 95% CI [-0.476 0.246]。在初始偏见上，任务之间没有显著的相关。

4.3.4　讨论

本研究中，我们的反应时与正确率（或 d prime）数据再次表明，在知觉加工中，对道德自我信息的加工具有明显的优势。不仅在匹配任务中，我们发现了道德自我相对于道德他人的优势（更快的反应时和更高的 d prime），而且在分类任务中也发现了同样的效应。

从知觉决策的过程来看，HDDM 分析再次表明，道德自我在知觉中的加工优势主要体现在信息累积的速率上：不管是在匹配任务中不是在分类任务中，道德自我的信息累积的速率均是最快的。与研究 8 相似的是，本研究中的 HDDM 也表明，在非匹配试次中，对自我相关的信息存在明显的反应初始偏见：被试更倾向于对自我相关的图形进行匹配的反应，这种反应表现在初始的偏见 z 上。

本研究与研究 8 一致地表明，道德自我的知觉加工优势主要是体现在扩散速率上。这种效应的模式与知觉决策中物理突显性类似，印证了之前研究的结果，即道德自我作为一种具有社会突显意义上的刺激，在当前的情境之中，具有像物理突显性一样的特点 (Sui et al., 2012)。

此外，本研究通过跨任务的分析表明，道德自我的知觉加工优势是具有跨任务的稳定性。这种加工优势的稳定可能表明，道德自我在信息加工中可能具有内隐的“锚定”作用，不随外界任务变化而变化，是一种固有的加工优势。接下来，本课题将进行进一步探讨道德自我的神经基础，加强对这种加工优势的理解。

第五章　道德自我加工优势的神经基础

5.1　本章引言

研究 1~9 已经分别从知觉层面上对道德自我的机制进行了探讨，这些研究表明，道德自我的内隐参照是道德信息在知觉中有优势的原因。而道德自我在知觉加工中的优势，主要是由于信息累积速率的优势，与物理突显性有类似的特点。为了进一步探索道德自我信息的加工如何在神经层面上实现的，本章将探讨道德自我优势的神经基础问题。

如前所述，心理自我的研究与道德两情境的研究均表明，默认网络参与到道德推理与自我信息的加工。由于道德自我在道德信息的加工中起到内隐参照的作用，因此道德自我可能会涉及到默认网络。

道德自我信息在知觉中的与物理刺激的突显性表现，道德自我的在认知加工中的优势，可能是由于默认网络与突显网络的交互作用完成。Goulden et al. (2014)发现，突显网络在 DMN 与执行网络之间的切换起着重要作用。在道德推理过程中，Chiong et al. (2013)发现，突显网络能够调节到默认网络的活动。在自我信息的加工中，Sui et al. (2013)发现，vmPFC 对与额顶中央执行网络的共同激活可能是自我信息加工产生优势的原因，而 vmPFC 正是默认网络的关键脑区。

由此可见，在道德信息与自我信息的加工中，vmPFC 均起到了重要的作用，而以前岛叶和背侧前扣带形成的突显网络也可能在信息加工的优势中起到关键作用。为了验证这种可能性，本研究采用神经成像的方法，初步探索道德自我的神经机制的问题。具体而言，本章将进行如何两部分研究。研究 10 通过探索大脑灰质体积与道德自我评估之间的关系，来探索道德自我加工的结构基础。研究 11 则采用神经成像的元分析方法，比较道德判断与自我参照共同激活的脑区。根据我们的道德自我内隐参照理论，在道德判断与自我参照加工中，均可能涉及到道德自我的内隐参照加工，因此两者共同激活的脑区可能是道德自我的功能相关脑区。

5.2 研究 10：道德自我与大脑结构的 VBM 分析

5.2.1 引言

为了探索道德自我相关的脑区，我们首先从结构的角度对其进行了探索，分析大脑结构与道德自我之间的关系。由于默认网络在道德判断中和自我参照中均起作用，因此，我们推断，默认网络的重要脑区 ACC/MPFC 以及 PCC 可能会在道德自我中起到重要的作用，这些脑区的结构特点可能与个人在道德自我上的差异有关。

然而，道德自我与 ACC/MPFC 的关系究竟是怎样的？有几种可能性：第一，认为自己越道德，脑区灰质越大。这表明人们有一种更强的自我增强的机制，能够保护自己。第二，如果人们对自己越客观，脑区灰质越大，表明 MPFC 起到的作用是对自己进行客观的认识。

为了检验这个假设，我们使用基于体素的形态学分析（Voxel-Based Morphometry，VBM）来分析道德自我与大脑灰质体积之间的关系。先前的研究已经表明，如果某项功能越强，负责该功能的大脑结构灰质体积可能会越大(Maguire et al., 2000)。

为了比较个体在道德领域中的自我增强效应，我们采用了社会心理学中常用的方法，即让被试评估自己与平均水平相比，大约处于什么水平。在本研究中，让被试从五个方面报告自己与平均水平相比处理什么水平(MacDonald, Saltzman, & Leary, 2003)。从道德得分中减去自我增强的成分，即得到道德自我的特异分数。

我们预期，道德自我增强得分将会与我们在研究中得到的元分析结果中的脑区存在相关，而且这种相关不受到性别、年龄、社会经济地位以及一般的自尊水平的影响。

5.2.2 方法

5.2.2.1 被试

本研究的数据来源于本实验进行的自我加工优势的神经基础与功能网络项目的一部分，共招募 250 名被试参与本次实验。所有被试均通过在线的广告招募。被试需要完成大脑扫描以及一系列的问卷和行为实验。在参加实验之前，所有被试均签署了知情同意书。所有数据收集自 2013 年~2015 年。

完成道德自我测量的包括 239 名被试（127 男性，年龄为 23.04 ± 2.57）；其中大脑结构有效被试量为 222 人（男 116 人，年龄：22.98 ± 2.58）。

5.2.2.2 道德自我的测量

在扫描前或扫描后，被试完成领域自评量表(MacDonald et al., 2003)，包括 5 个问题，分别让被试报告自己在五个领域上与平均水平进行相比：能力、身体吸引力、物质财富、社交能力和道德（见附录 A）。被试从 1-12 来评估自己相对于平均水平而言，在这 5 个维度上的位置。其中，关于道德自我评分反映了被试对自己在道德水平上的自我认识。在本研究中，该问卷的信度为：$\alpha = 0.76$，95% CI[0.72 0.81]。

为了克服单一测量的偏差，除了领域自评量表之处，还要要求被试进行自我特质词评估。这些特质词包括 36 形容词(Wang et al., 2016)，包括 18 个积极词汇和 18 个消极词汇（见附录 B）。

由于根据特质词并非根据社会知觉的三个维度而选定，因此，我们采用了两种分析方法，来分离出道德、能力与社会能力三个维度的得分。首先，我们招募一批独立的被试（N = 82，29 名男性被试，年龄：18.7 ± 1.1 岁）对这些特质词进行判断。要求他们判断每个词适合于描述人的哪个方面的特点，能力、道德、还是社交能力？如果一个词汇适合于多个方面，则选择“无法区分”；如果一个词汇不适合于这三个方面中的任何一个，则选择“其他”方面。假如某个词汇被 60%及以上的被试选择为属于三个维度中的一个，将其归属于该维度。根据词汇选择归属，我们对词汇对进行分维度计分。

其次，我们采用因子分析(Kitayama et al., 2014; Wang, Peng, Chechlacz, Humphreys, & Sui, 2017)的方法，进一步对三个维度进行分析，使用因子得分作为该维度的得分。具体而言，我们将上述的特质词与被试自我报告五个方面中的道德、社交能力和能力三个项目的得分一起，共同进行因子分析。通过回归的方法来计算被试综合的道德自我判断得分。

为了控制一般自尊的影响，加入了被试的自尊得分作为控制变量(Rosenberg, 1965)（见附录 C）。该问卷包括 10 个项目，被试需要根据他们自己情况，评估每个维度是否符合自己的情况，从 1-4 进行打分，其中 1 表示完全不符合，4 表示完全符合。本研究中自尊问卷的信度为$\alpha = 0.85$，95% CI [0.82 0.87]。

5.3.2.3 图像成像扫描

所有 MRI 图像均在 Philips Achieva 3.0T TX 系统进行采集，采用 SENSE 8 通

道头部线圈。T1 结构像包括 160 个 1 mm 厚的矢状切片组成，TR = 8.2 ms，TE = 3.8 ms，矩阵大小为 256 体素 * 256 体素，每个体素的体积为 0.938 mm * 0.938 mm * 1 mm ，flip angle 为 8 度。采用 Sensitive encoding (SENSE)平行扫描成像技术(Pruessmann, Weiger, Scheidegger, & Boesiger, 1999)以减少成像时间。SENSE 加速技术在前-后方向的加速因子为 2，右-左方向为 1.5。

5.2.2.4 脑成像数据的预处理

使用 SPM(SPM8; Wellcome Department of Cognitive Neurology, London, United Kingdom; www.fil.ion.ucl.ac.uk/spm)以及 DPABI 软件包(Yan, Wang, Zuo, & Zang, 2016; Yan & Zang, 2010)进行预处理。图像预处理流程具体如下。

（1）使用 SPM8 对每个被试的 T1 图像均进行检查，调整其朝向与朝向点位置，使其尽量与 MNI 模板保持一致，提高配准效果

（2）使用 New Segmentation 从 T1 图像中分离灰质、白质、脑脊液、骨骼、头皮等结构，并将灰质和白质图像转换成适用于 DARTEL 工具包的格式；

（3）使用 DARTEL(Ashburner, 2007)模块，以分割出灰质图像，并建立基于样本的灰质模板，与 MNI 模板配准，并计算出各个被试的变形流场；

（4）使用 Normalize to MNI 模块，以 DARTEL 模块和变形流场为基础，将被试的 T1 灰质图像配准至 MNI 空间；

（5）使用[4mm 4mm 4mm]的高斯平滑滤波器对数据进行平滑处理；配准方式采用调制（modulation），使结果图像中每个像素数值代表该位置的灰质体积。

最后，将图像重建为 1.5mm × 1.5mm × 1.5mm 分辨率的大脑图像。

5.3.2.5 统计分析

统计分析均在 SPM8 中完成。本研究中进行感兴趣区分析和全脑分析。

感兴趣区分析（ROI analysis）　基于先验假设，以 Hu et al. (2016)对心理自我神经成像元分析的结果为基础，定义了三个关键的感兴趣区。第一个 ROI 为心理自我激活最强烈的脑内侧前额叶及前扣带回（mPFC/ACC，MNI 坐标[-1 47 6]），该脑区被认为是自我加工的核心区域；第二个 ROI 为心理自我与身体自我共同激活的脑区左侧前岛叶（left anterior insula，MNI 坐标[-35 12 -6]）；第三个 ROI 也是心理自我与身体自我共同激活的脑区前扣带回（ACC，MNI 坐标：[1 35 19]）。具体方法如下：（1）使用 MarsBar 工具包(Brett, Anton, Valabregue, & Poline, 2002)，根据 Hu et al. (2016)中报告的 cluster 大小，分别定义了三个半径为 12mm, 7mm 和 6mm 的球形 ROI；（2）将 ROI mask 重新切割成为 1.5mm × 1.5mm × 1.5mm 分辨率，以与后面 T1 功能像保持一致。

随后使用预处理后的灰质图像（smwc*），分别在 3 个 ROI 内区域内进行基

于体素的 GLM 分析，因变量为像素对应的灰质体积，预测变量为被试的道德自我评分，控制变量包括被试的性别、年龄以及基于解剖图像计算出来的颅内总体积（total incranial volume, TIV），以控制头部大小的影响。此外，根据 VBM 根据的惯例，使用 0.1 的绝对阈限 mask，以排除灰质－白质交界处的噪音。基于较强的先验假设，体素水平采用了较为宽松的阈限 $p_{uc} < 0.05$。多重比较通过 Resting-State fMRI Data Analysis Toolkit (REST)中的 AlphaSim 算法控制；5000 次 Monte-Carlo 模拟（参数：rmm=2.3mm，FWHM = [7.6mm 7.5mm 7.4mm]）显示，采用 k > 375 的簇水平阈限，可以得到 $p < 0.05$ 的校正后水平。此外，k > 291 的簇水平阈限可以得到 $p < 0.10$ 的校正后水平；通过了簇水平 $p < 0.05$ 未校正的区域可以被看作是边缘显著。为了控制 nonisotropic 平滑中的问题，在计算簇水平的值时，使用 Nonstationary extent correction。

全脑分析（whole-brain analysis） 为了探索其他脑区与道德自我得分的相关性，我们在全脑范围内进行与 ROI 相似的 GLM 分析，因变量为像素对应的灰质体积，预测变量为被试的道德自我评分，控制变量包括被试的性别、年龄、自尊得分以及基于解剖图像计算出来的颅内总体积。并使用 0.1 绝对阈限的 mask。体素水平阈限设定为 $p_{uc} < 0.05$；根据本研究的样本量（222），使用这一阈限水平，可以以 80%的统计效力检验 r = 0.24 的效应量。同样使用 AlphaSim 控制多重比较问题；5000 次 Monte-Carlo 模拟（参数：rmm=2.3mm，FWHM = [7.6mm 7.5mm 7.4mm]）显示，采用 k > 375 的簇水平阈限，可以得到 $p < 0.05$ 的校正后水平。此外，k > 291 的簇水平阈限可以得到 $p < 0.10$ 的校正后水平；通过了簇水平 $p < 0.05$ 未校正的区域可以被看作是边缘显著。在计算簇水平的 p 值时，进行了 Nonstationary extent 校正。

对每个显著的脑区均创建了散点图。图中显著的是道德自我得分与 cluster 的峰值之间的相关系数以及其 95%的置信区间。

5. 2. 3　结果

5.2.3.1　道德自我评估得分的计算

根据我们结果显示，有 36 个词中有 16 个属于道德、能力和社交能力三个维度（表 5.1 全部词汇见附录 B）。

表 5.1 特质词在能力、社交能力和道德维度的归属情况(%)

特质词	能力	社交	道德
自私的			81
无私的			81
外向的		88.9	
开朗的		85.7	
热心的		85.7	
内向的		79.4	
沉默的		76.2	
羞涩的		74.6	
主动的	20.6	71.4	
冷漠的		68.3	20.6
安静的		65.1	
大方的		63.5	14.3
健谈的	42.9	54.0	
聪明的	88.9		
自律的	71.4		
粗心的	69.8		
愚蠢的	79.4		

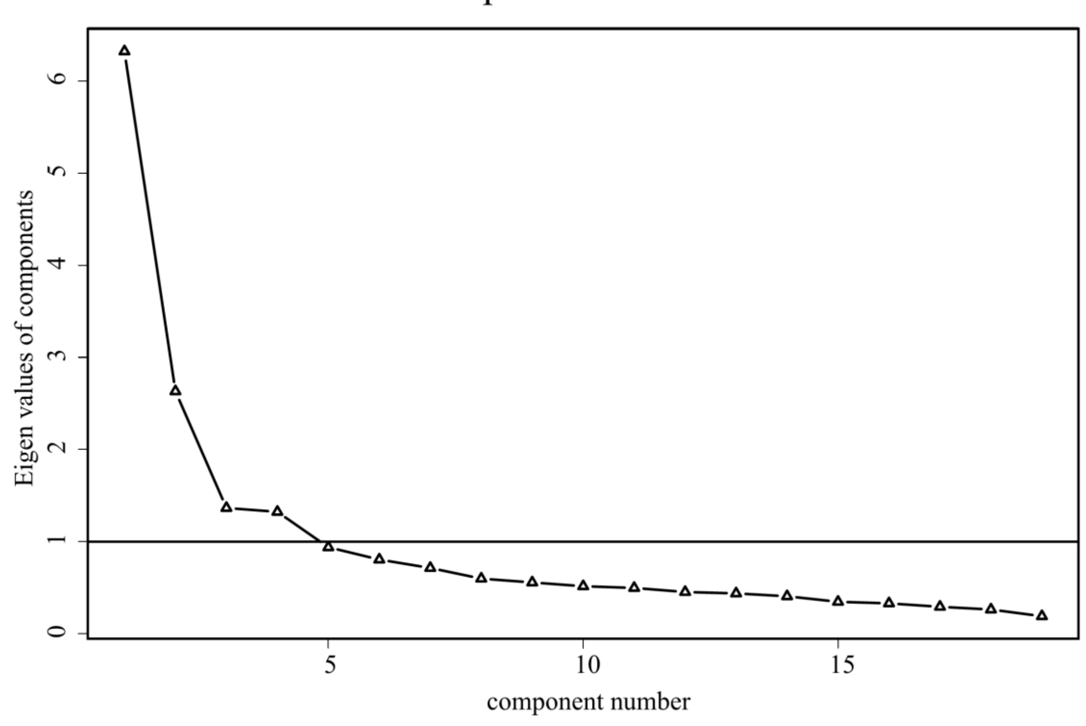


图 5.1　领域自我评估与特质词汇自我评估。

表 5.2 自我评分词汇与领域自尊的因子负荷

	外向性	道德	责任
内向的	0.83		
安静的	0.72		
沉默的	0.71		
羞涩的	0.65		
外向的	-0.63		
无私的		0.75	
道德领域		0.63	
自私的		-0.71	
聪明的			0.80
能力领域			0.52

注：加粗部分为领域自我评估内容。

将特质词与领域自主量表中的道德、社交能力和能力三个维度的评分进行因子分析。数据的 KMO 值为 0.87，大于通常的 0.6 标准，适用进行因子分析。如下

图所示，尽管有 4 个因子的性值大于 1，但是在第三个因子处有明显的拐点，因此按 3 个因子提取。采用 the principal-axis-factoring method and oblimin rotation with Kaiser normalization 提取出因子，计算各条目的载荷并删除多重载荷以及载荷小于 0.4 的条目 9 条，最后得到 10 个条目 3 因子结构。在这个模型中，因子 1 代表的是外向性；因子 2 代表的是道德水平，因子 3 代表的是能力。三个因子累积解释了变异的 54%。

这个模型之下，其中内向性 5 个项目的 alpha 系数为 0.83，道德维度的 3 个项目的 alpha 系数为 0.72，责任维度的 2 个项目为 0.63（表 5.2）。采用道德维度因子分析中的因子得分作为被试道德自我评估测量指标。该模型与由被试独立进行评估的结果一致，表明这些词在维度归属上是稳定的结构。

5.3.3.2 道德领域自我增强的优越性

为了检验个体在不同领域的自我增强水平是否有差异，我们对领域自评结果进行了方差分析，以自评的领域作为被试内变量，性别作为被试间变量进行了混合方差分析。结果表明，性别与自尊领域的交互作用显著：$F(4,956) = 3.36, p = 0.0096, \eta_G^2 = 0.0067$; 性别的主效应不显著：$F(4,956) = 1.48, p = 0.23, \eta_G^2 = 0.0032$，道德领域的主效应显著：$F(4,956) = 26.11, p < 0.0001, \eta_G^2 = 0.05$。

对道德领域主效应的事后比较表明：道德领域（9.31 ± 1.87）的得分高于能力（8.4 ± 1.77）、身体吸引力（7.18 ± 2.12）、社交能力（7.13 ± 2.26）和财富（5.56 ± 1.97）。这表明，道德领域确实存在着自我增强的特异性。

虽然性别不同的自评领域具有交互作用，由于本研究的主要关注点在道德领域，因此仅对男女在道德领域上的自评是否有差异。t 检验结果表明，在道德领域上，男女没有显著差异：$t(222) = -0.101, p = 0.92$，其中，男性为 9.32 ± 1.72；女性为 9.30 ± 2.03。

为了计算自我道德优越性指标，我们使用道德上的得分减去存在自我增强效应领域的平均得分。首先，为验证自我增强效应，我们将每个维度的得分与均值 6 进行比较，检验每个维度上的得分均值是否显著大于 6。结果表示，除了在财富这个维度外，其他四个领域上均存在自我增强的效应。道德领域，$t(240) = 27.482, p < 0.0001$;能力领域，$t(240) = 21.13, p < 0.0001$; 社交能力领域：$t(240) = 7.77, p < 0.0001$; 身体吸引力，$t(240) = 8.654$，$p < 0.0001$。在财富领域，平均得分显著小于 $t(240) = -3.5, p = 0.00056$。因此，我们使用道德领域的得分减去能力、社交能力和身体吸引力三个领域的平均得分，得到道德自我优越性得分（Moral Self Superiority Index，MSSI）。随后我们比较了男性在道德自我优越性得分上的差异，结果表明，

男女之间没有显著的差异 $t(217.86) = -1.3539$, $p = 0.1772$。

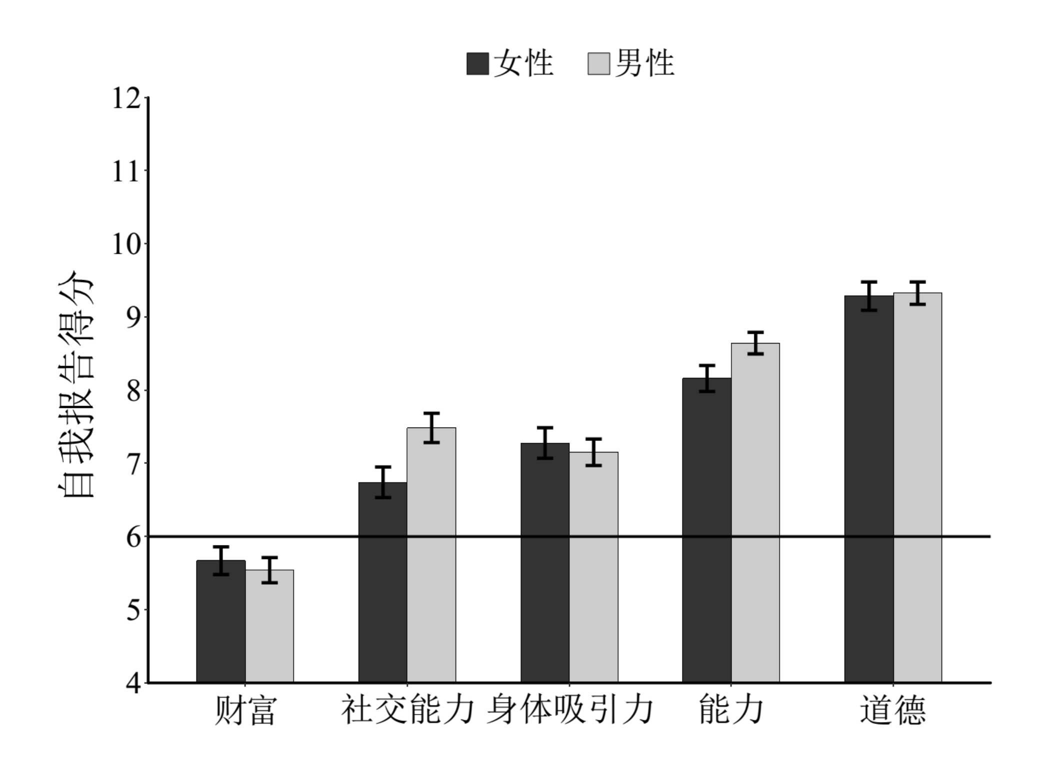


图 5.2　被试在各个领域的自我报告得分。

5.3.3.3 VBM 结果：ROI 分析

ROI 的分析结果表明，在三个 ROI 范围内，未发现某个脑区的灰质体积与道德自我评分有相关。采用更加宽松的标准，我们在左侧岛叶的感兴趣区内发现两个簇与道德自我评分呈正相关：峰值坐标[-32 10 -1]，k = 16；峰值坐标[-38 18 -10]，k = 28。

5.3.3.3 VBM 结果：全脑分析

全脑分析也未能发现通过多重比较校正的脑区与道德自评得分呈正相关或者负相关。

使用更加宽松的标准进行探索性分析表明（体素水平 $p < 0.001$），有两个簇与被试的道德自评得分呈正相关：一个是在前楔叶区域发现一个体积为 128 体素的簇，峰值坐标为[53 -65 23]；另一个的簇大小为 124 体素，峰值坐标为[54 -48 -24]。

5.2.4　讨论

本研究的行为结果表明，被试在行为水平上存在显著的自我增强效应，表现为在身体吸引力、能力和社交能力的自我评估上，表现出高于平均水平的评分。本研究的结果证实了先前研究的结论，即在道德领域的自我增强效应最为显著，而且在道德领域的自我增强效应没有性别的差异。而在 VBM 上，我们未发现与道德自我评估有显著相关的大脑区域。

VBM 分析的结果未发现与道德自我评分存在显著相关的脑区。虽然使用更加宽松的标准，我们发现有小簇体积的大脑区域与道德自我评分存在着正相关。但是由于全脑分析中存在着严重的多重比较问题(Bennett, Miller, & Wolford, 2009)，因此这个激活是否具有意义还有待于进一步讨论。

同样，在我们定义的三个 ROI 区域内，也只在较为宽松的标准（voxel level, $p < 0.05$ uncorrected）下发现有小的簇激活。

结构成像的阴性结果可能有多种原因。一种可能是道德自我评估本身属于高级的认知功能，因此可能没有单个的脑区与道德自我评估直接相关。第二种可能性是由于本研究中的道德自我评分是否能够真正代表被试的道德自我评估。由于本研究属于数据驱动，采用已经获得的数据进行分析，因此从实验设计角度来讲，可能统计的效力不足。例如，本研究中的道德自我评估抽取自被试对 36 个特质词的自我评估，而这些词本身并非用来道德研究，因此最终通过统计分析被归入道德维度的词汇较少。

虽然当前研究中的 VBM 分析未能发现道德自我与特定脑区结构上的相关，但是我们发现，被试报告的结果中，道德自我增强效应本身却非常稳定，这将为以后进一步探索神经机制提供很好的基础(Krakauer, Ghazanfar, Gomez-Marin, MacIver, & Poeppel, 2017)。

5.3　研究 11：道德自我神经机制的 ALE 元分析

5.3.1 引言

研究 10 从大脑结构的角度讨论了道德自我相关的脑区。为了进一步探索道德自我评估的神经机制，在本研究中，采用基于坐标的元分析方法(Wager et al., 2007; 胡传鹏 et al., 2015)，从功能的角度来探索加工道德自我相关的神经基础。

采用元分析方法的主要原因有两个。从理论假设的角度来讲，如前所述，道德判断与自我参照加工在认知机制上，可能均涉及到内隐的道德自我参照加工，在神经机制上均激活了默认网络相关的脑区，因此默认网络可能是道德自我的加工的神经基础。

从研究方法的角度来讲，元分析方法的优势是可以将采用不同手段的研究进行比较，尤其是当这些方法很难在同一个实验中完成时(Wager et al., 2007; 胡传鹏 et al., 2016)，这种分析比较的结果，可以为提出假设或者验证假设提供证据。

为了验证道德自我的内隐参照加工假设，在本研究中，采用神经成像元分析方法来探索道德自我的神经机制，然后以此作为今后神经成像研究的基础，将有利于提出具体的研究的假设。

5.3.2　方法

5.3.2.1　文献搜索与挑选

通过 PubMed 数据库和 Google Scholar(到 2016 年 12 月)进行系统搜索获得。具体而言，对于心理自我的文献采用如下的关键词：“self-referential”和“self-reference”，这两个关键词与“fMRI”和“PET”使用“AND”链接来搜索相关的文献。对于道德判断相关的文献，采用如下关键词：“moral judgment”、“morality”、“ethic judgment”和“social norm”，同样，这些关键词也与“fMRI”和“PET”使用“AND”链接对相关文献进行搜索。为了更加完整地包括所有的文献，我们也参考了近期发表的元分析(Araujo, Kaplan, & Damasio, 2013)和综述。进入元分析的文献需要满足如下的标准：

（1）仅报告了全脑分析结果的研究被包括到元分析之中，只包括部分大脑或者使用感兴趣区（region-of-interest, ROI）分析的研究被排除；

（2）使用标准的大脑空间报告空间坐标结果（Talairach 或者 MNI）的研究被

包括进来。为了解决两个标准空间坐标之间的转化问题，使用 Talairach 空间坐标的研究结果被转化成为 MNI 的空间坐标结果，使用 Lancaster 转换算法(Lancaster et al., 2007)；

（3）使用健康被试的研究被包括到元分析之中，而神经疾病或者其他精神异常的被试的报告则被排除；仅年轻和中年成年为被试的研究被包括进来（18－59 岁），从而避免年龄变化导致的自我相关加工的异常(Leshikar & Duarte, 2014)。

（4）如果同一组数据被报告在不中的文献之中，则我们仅选择其中一个进入元分析；

（5）仅包括了使用视觉方式呈现刺激的研究被包括进来，使用其他的感觉通道进行刺激呈现的研究被排除。

使用如上标准，在心理自我加工方面，最终得到 37 篇文献进入元分析，包括 41 个实验，733 个被试，215 个坐标点。在道德判断方面，最终得到 62 篇文献，包括 63 个实验，1430 个被试，455 个坐标点。与先前的研究相似(Murray, Schaer, & Debbané, 2012)，在对心理自我的元分析中，我们选择了 self -referential > other-referential 条件的坐标点作为心理自我的结果。在道德判断的元分析之中，我们选择 moral judgment > non-moral judgment 这种比较的结果作为道德特异加工的脑区结果（见表 5.3）。

表 5.3　包括在元分析之中的实验信息

Articles	Model	Subjects	Mean age	Task
Benoit, et al., 2010	fMRI	13 (8M)	24.2	Self-referential
Cikara, et al., 2014	fMRI	21 (10M)	23.1	Self-referential
D'Argembeau, et al., 2007	fMRI	17 (6M)	23	Self-referential
Fossati, et al., 2003	fMRI	10 (3M)	25.8	Self-referential
Gutchess, et al. 2007	fMRI	19 (9M)	23.1	Self-referential
Gutchess, et al. 2010	fMRI	17 (9M)	23.2	Self-referential
Han, et al., 2008	fMRI	14 (8M)	22.5	Self-referential
	fMRI	14 (6M)	23.6	Self-referential
Han, et al., 2010	fMRI	14 (7M)	25.4	Self-referential
Heatherton, et al., 2006	fMRI	30 (14M)	24	Self-referential
Jenkins, et al., 2011	fMRI	15 (5M)	21.2	Self-referential
Kelley, et al., 2002	fMRI	21 (12M)	20	Self-referential
Lombardo, etal., 2010	fMRI	33 (33M)	27.9	Self-referential
Lou, et al., 2004	PET	13 (6M)	27	Self-referential
Ma & Han, 2011	fMRI	23 (11M)	22	Self-referential
Ma et al., 2014a	fMRI	30 (15M)	22.7	Self-referential
	fMRI	30 (15M)	25	Self-referential
McAdams, et al., 2012	fMRI	18 (0M)	18~45	Self-referential
Modinos, et al.,2009	fMRI	16 (10M)	20.8	Self-referential
Modinos, et al.,2011	fMRI	18 (10M)	20.8	Self-referential
Moran, et al., 2011	fMRI	21 (8M)	24.3	Self-referential
Ochsner, et al.,2005	fMRI	16 (7M)	22.95	Self-referential
Pauly, et al., 2014	fMRI	13 (7M)	34.46	Self-referential
Pfeifer, et al., 2007	fMRI	12 (6M)	26.1	Self-referential
Powell, et al., 2010	fMRI	13 (4M)	20	Self-referential
Sarsam, et al., 2013	fMRI	14 (6M)	26.4	Self-referential
Schmitz, et al., 2004	fMRI	18 (10M)	24	Self-referential
Sul, et al., 2011	fMRI	19 (9M)	24	Self-referential

表 5.3（续）　包括在元分析之中的实验信息

Articles	Model	Subjects	Mean age	Task
van Buuren, et al., 2010	fMRI	19 (8M)	21.5	Self-referential
van Buuren, et al., 2012	fMRI	18 (7M)	27.5	Self-referential
Vanderwal, et al., 2008	fMRI	17 (10M)	21.5	Self-referential
Wang, et al, 2012	fMRI	32 (16M)	22.7	Self-referential
Wu et al., 2010a	fMRI	16 (8M)	20.1	Self-referential
	fMRI	15 (8M)	19.8	Self-referential
J. Yang, et al., 2012	fMRI	17 (9M)	22.4	Self-referential
W. Yang, et al., 2013	fMRI	25 (11M)	22.5	Self-referential
Yoshimura, et al, 2009	fMRI	15 (7M)	23.3	Self-referential
Zhang, et al., 2006	fMRI	7 (4M)	20.4	Self-referential
Y. Zhu, et al., 2007	fMRI	13 (8M)	21.5	Self-referential
	fMRI	13 (8M)	23.1	Self-referential
L. Zhu, et al., 2012	fMRI	14 (4M)	22.6	Self-referential
Avram et al., 2013	fMRI	16 (7M)	28.5	Moral judgment
Bahnemann, et al, 2010	fMRI	25 (22M)	26	Moral judgment
Caldwell, et al, 2015	fMRI	36 (36M)	29.19	Moral judgment
Chakroff, et al., 2016	fMRI	23 (20M)	27	Moral judgment
Cherry, et al., 2015	fMRI	23 (NA)	29.65	Moral judgment
Chiong, 2013	fMRI	16 (NA)	66	Moral judgment
Christensen, 2014	fMRI	24 (NA)	23.51	Moral judgment
Cikara, 2010	fMRI	13 (6,20.7)	20.7	Moral judgment
Cloutier, 2012	fMRI	19 (8M)	24.2	Moral judgment
Cloutier, 2014	fMRI	20 (20M)	24.3	Moral judgment
Cushman	fMRI	35 (15M)	18-28	Moral judgment
Decety, 2011	fMRI	22 (22M)	19-26	Moral judgment
Denke, 2014	fMRI	17 (9M)	25	Moral judgment
Englander, 2012	fMRI	10 (3M)	20-27	Moral judgment
FeldmanHall, 2014	fMRI	38 (18M)	24.6	Moral judgment

表 5.3（续）　包括在元分析之中的实验信息

Articles	Model	Subjects	Mean age	Task
Greene, 2004	fMRI	41 (24M)	NA	Moral judgment
Greene, 2001a	fMRI	9 (5M)	NA	Moral judgment
Greene, 2009	fMRI	28 (17M)	24	Decision
Han, 2014	fMRI	16 (8M)	28	Moral judgment
Harda, 2009	fMRI	18 (8M)	25.3	Moral judgment
Harenski, 2006	fMRI	10 (0M)	19-29	Picture viewing
Harenski, 2008	fMRI	28 (14M)	25.8	Moral judgment
Harenski, 2010a	fMRI	14 (0M)	24.5	Implicit task
Harenski, 2010b	fMRI	32 (32M)	33.3	Moral judgment
Harenski, 2012	fMRI	36 (36M)	27.3	Moral judgment
Harrison,2012	fMRI	73 (42M)	32.6	Moral judgment
Hayashi, 2014	fMRI	37 (18M)	21.2	Moral judgment
Heekeren, 2003	fMRI	8 (7M)	27.8	Moral judgment
Heekeren, 2005	fMRI	12 (10M)	25.75	Moral judgment
Hutcherson, 2015	fMRI	26 (14M)	27.7	Moral judgment
Kahane, 2012	fMRI	16 (7M)	29.25	Moral judgment
Li, 2015	fMRI	24 (12M)	20.42	Moral judgment
Luo, 2008	fMRI	20 (9M)	20-36	IAT
Majdandzic, 2012	fMRI	40 (40M)	19.2	Moral judgment
Mende-Siedlecki, 2013	fMRI	23 (10M)	18-31	Moral judgment
Molenberghs, 2016	fMRI	48 (24M)	22.2	video viewing
Moll, 2001	fMRI	10 (6M)	24-43	Moral judgment
Moll, 2002a_NI	fMRI	7 (3M)	30.3	Moral judgment
Moll, 2002b_JN	fMRI	7 (5M)	27-31	Picture viewing
Ngo, 2015	fMRI	16 (8M)	24	Intention judgment
Nichelli, 1995	PET	9 (NA)	61.5	Moral judgment
Parkinson, 2011	fMRI	38 (NA)	NA	Moral judgment
Preckel, 2015a	fMRI	24 (23M)	24.13	Moral judgment

表 5.3（续）　包括在元分析之中的实验信息

Articles	Model	Subjects	Mean age	Task
Prehn, 2008	fMRI	23 (0M)	25.17	Moral judgment
Pujol, 2012a	fMRI	22 (22M)	40.6	Moral judgment
Pujol, 2012b	fMRI	22 (22M)	39.8	Moral judgment
Reniers,2012	fMRI	24 (24M)	18-30	Moral judgment
Robertson, 2007	fMRI	16 (16M)	35	Moral judgment
Schaich Borg, 2006	fMRI	24 (13M)	22.9	Moral judgment
Schaich Borg, 2008	fMRI	15 (15M)	25	Memory
Schaich Borg, 2011a	fMRI	26 (26M)	32	Moral judgment
Schleim, 2011	fMRI	40 (22M)	31.05	Moral judgment
Schneider, 2013a	fMRI	28 (14M)	31.29	Moral judgment
Schneider, 2013b	fMRI	28 (14M)asd	31.39	Moral judgment
Shenhav, 2010	fMRI	34 (17M)	24.3	Moral judgment
Shenhav, 2014	fMRI	35 (16M)	22.9	Moral judgment
Smith, 2015	fMRI	30 (5M)	26	Moral judgment
Sommer, 2010	fMRI	12 (5M)	24.5	Moral judgment
Sommer, 2014a	fMRI	16 (8M)	24.9	Moral judgment
Sommer, 2014b	fMRI	16 (7M)	**15**	Moral judgment
Takhashi, 2008	fMRI	15 (NA)	21.1	Reading task
Tsukiura, 2011	fMRI	20 (0M)	23.4	Moral judgment
Verdejo-Garcia, 2014a	fMRI	14 (1M)	35.1	Moral judgment
Wang, 2015a	fMRI	28 (14M)	22	Moral judgment
Wang, 2015b	fMRI	22 (10M)	21	Gender judgment
Yamada, 2012	fMRI	22 (11M)	21.5	Moral judgment
Yoder, 2014	fMRI	40 (19M)	21	Moral judgment
Young, 2007	fMRI	10 (4M)	18-22	Moral judgment

5. 3. 2. 2 元分析方法

所有的元分析均使用 ALE（activation likelihood estimation）算法(Eickhoff et al., 2009; Turkeltaub, Eden, Jones, & Zeffiro, 2002; Turkeltaub et al., 2012)，使用 Matlab 代码实现。ALE 算法的目标是确定那些显示出跨实验一致的坐标，这些坐标点激活的可能性高于随机的空间分布的可能性。ALE 的核心的原理是将文献中报告的坐标点当作一个以报告坐标点中中心的 3D 高斯概率分布，从而来模拟每个坐标点上空间分布的不确定性。而这个 3D 高斯分布的半径，则由实证数据中被试间和大脑模板之间的变异来决定，因为这两个变异反应了最主要的变异。更重要的是，算法对所有不同被试数目的实验进行加权，因为样本量大的研究通常有更加稳定的结果，因此对真实的激活有更好的估计，其 3D 高斯分布的半径更小(Eickhoff et al., 2009)。

随后，所有坐标点的激活概率均在体素水平进行联合，形成一个模型化的激活图（modeled activation, MA）。对每个坐标点的 MA 图的联合，得到一个激活可能性的结果图，其中每个体素上均有激活可能性的估计值(activation likelihood estimation, ALE)的量化值。为了将真正的跨实验的激活与随机的噪音区分开来，ALE 得分将与一个通过 permutation 得到的零分布进行比较，从而得到真正与实验相关的激活的收敛。这个比较可以得到随机效应的推断，从而得到高于随机水平的跨实验的激活，多面角是某个特定实验激活。从算法上讲，得到这个零分布是将所有的坐标点在大脑空间内进行随机分布之后，计算出每个体素在这种随机分布条件下的激活可能性，然后将这种随机分布的坐标点的 MA 图进行联合，得到每个体素上的激活可能性。如此重复 5000 次，记录每个体素上在 5000 次中最大的激活可能性，形成了零分布。然后些许每个体素上真实的激活可能性与零分布中的激活可能性，得到一个关于真实的激活可能性是否高于零分布下激活可能怀的 p 值。为了校正多重比较的问题，采用 Cluster-level $p < 0.05$ 校正方法(Eickhoff, Bzdok, Laird, Kurth, & Fox, 2012)。我们报告了所有显著的 cluster，包括其体素的数量、中点位置、大脑区域以及该区域的峰值 Z 值。

5. 3. 2. 3 对单个元分析的联合分布以及对比分析

心理自我与道德判断之间对比分析（contrast）和联合分析（conjunction）与单个神经成像的元分析采用类似的思路。先分别计算出两种条件下的 ALE 图，然后在体素水平上对两种条件下的 ALE 图相减，得到两种条件差别的 ALE 图。随后，也使用随机 permutation 的方法来得到两者 ALE 差异图的零分布：将两种条件下所

有坐标点混合之后，按照两种条件下坐标点的个体数进行随机分组，得到两组随机的坐标点，然后再对这两组随机的坐标点计算 ALE 图以及其差值，同样对两组随机坐标点的 ALE 图进行相减，得到随机的 ALE 差异值，记录每个体素上的差异值。重复 25000 次之后，得到一个 ALE 差异值的零分布图，然后比较真实的差异值与差异值零分布图，从而得到真正的差异值。对于差异图的多重比较校正，我们采用了 $p < 0.01$ (FDR 校正)。

5.3.2.2.3 数据的可视化

所有元分析得到的大脑区域对应的生理解剖位置均使用人脑的细胞构筑可能性图（probabilistic cytoarchitectonic maps）进行定位(Eickhoff, Heim, Zilles, & Amunts, 2006; Eickhoff et al., 2007; Eickhoff et al., 2005)。BrainNet Viewer(Xia, Wang, & He, 2013)被用来可视化元分析的结果。

5.3.3　结果

5.3.3.1　心理自我的元分析结果

在自我参照条件下比他人参照条件下一致更加激活的脑区见下图和下表。与以往研究相似的是，皮质中线结构（cortical midline structure, CMS）非常稳定地出现在激活之中，包括双侧的前扣带回/内侧前额叶（ACC/medial frontal gyrus）。其他的大脑结构如左侧的额中回（middle frontal gyrus）也在自我参照条件下更加激活（见表 5.4；图 5.3）。

表 5.4　单个元分析的结果

脑区	体积 (voxels)	脑区中心位置 x	y	z	Z-值峰值	Z-值峰值坐标 x	y	z	解剖位置
自我参照条件 > 他人参照条件									
1	1798	-1	47	6	5.94	-4	36	16	ACC
					5.85	-4	38	2	ACC
					5.84	-4	60	2	L mid orbital gyrus
					5.68	-2	42	10	L ACC
					5.25	-10	50	0	L mid orbital gyrus
					5.10	8	42	24	R ACC
					4.65	8	60	-6	R mid orbital gyrus
2	257	-25	48	29	4.62	-30	48	18	L middle frontal gyrus
					4.37	-22	40	42	L superior frontal gyrus
					4.15	-22	56	30	L superior frontal gyrus
					3.95	-24	52	24	L middle frontal gyrus
道德判断条件 > 其他判断条件									
1	723	-4	52	27	5.35	-4	52	32	L superior medial gyrus
					5.30	-6	54	12	L superior medial gyrus
					4.58	4	58	14	R superior medial gyrus
					4.02	-10	44	44	L superior medial gyrus
2	591	-49	-62	25	7.65	-48	-62	24	L middle temporal gyrus
3	433	0	-56	32	7.82	0	-58	32	L PCC
4	296	53	-57	27	5.02	54	-54	28	R superior temporal gyrus
					3.87	56	-60	22	R Middle temporal gyrus
5	128	-1	54	-10	3.99	-4	54	-12	L rectal gyrus
					3.90	0	54	-6	L mid orbital gyrus
					3.18	-10	50	-8	L mid orbital gyrus
6	93	-39	24	-22	5.13	-38	24	-22	L temporal pole

所有的解剖位置均通过 SPM Anatomy 工具进行定位.

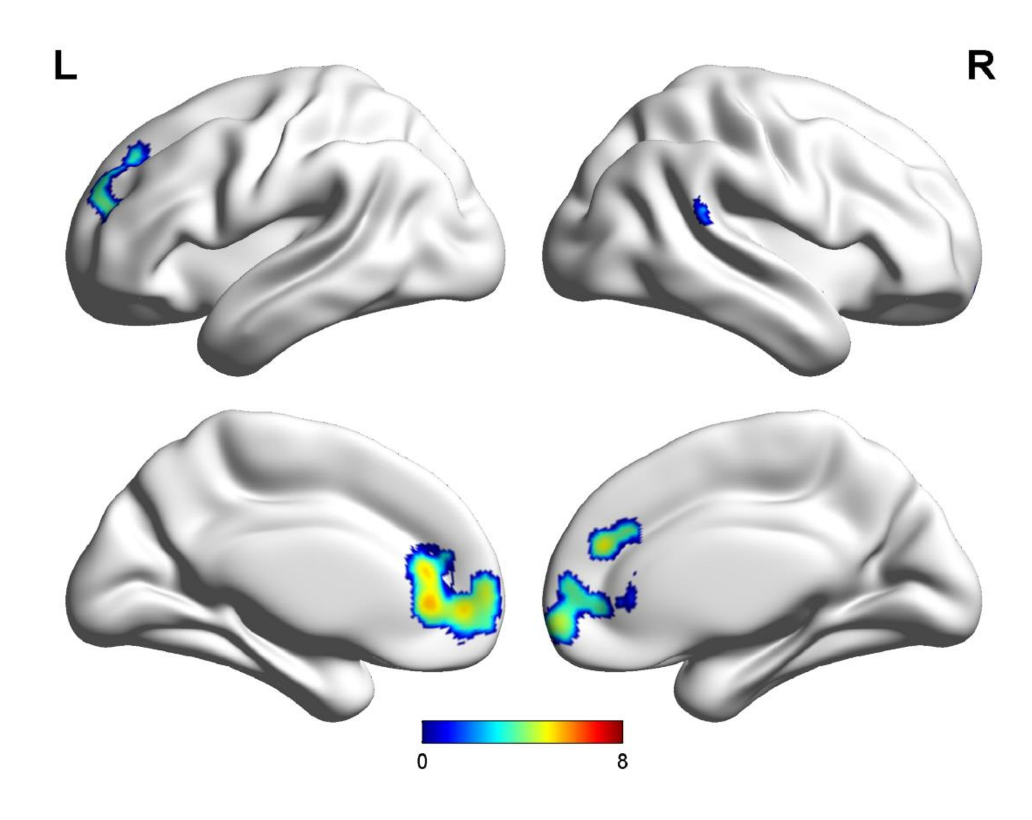


图 5.3　对心理自我神经成像研究元分析的结果。

5.5.3.2 道德判断的元分析结果

以道德判断 > 其他判断的结果作为输入，我们发现，道德判断任务一致激活了双侧的内上额叶（superior medial frontal gyrus）、左侧 PCC、双侧的颞中回（middle temporal gyrus）、右侧颞上回（right superior temporal gyrus）、左侧眶额皮层（left mid orbital gyrus）（表 5.4；图 5.4）。

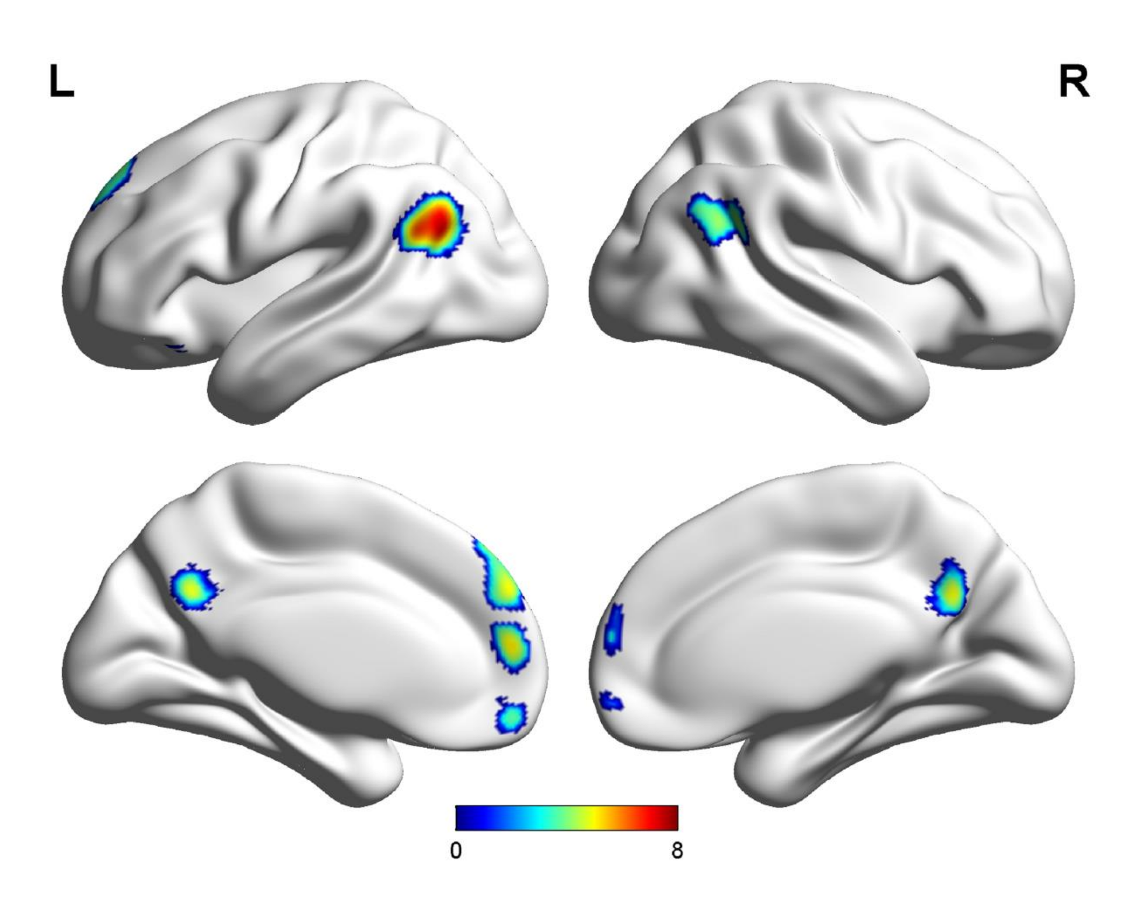


图 5.4　对道德判断神经成像元分析的结果。

5.4.3.3 心理自我与道德判断激活脑区的异同

为了研究心理自我与道德判断在神经机制上的异同，我们比较了两个 ALE 元分析的结果。对比分析（contrast）表明，双侧 ACC/mid orbital gyrus、左侧 middle frontal gyrus、右侧 superior temporal gyrus 更多在自我加工中出现；而双侧的 middle temporal gyrus、PCC、左侧 temporal pole 和左侧的 superior medial frontal gyrus 则更多在道德判断中出现（表 5.5；图 5.5）。

对心理自我与道德判断的元分析结果的联合分析表明，左侧 mid orbital gyrus 和左侧的 Superior medial gyrus 在两种条件下均激活（表 5.5；图 5.6）。

表 5.5　心理自我与道德判断元分析结果的对比与联合分析

脑区	体积	脑区中心位置			Z-值峰值	Z-值峰值坐标			解剖位置
	(voxels)	x	y	z		x	y	z	
自我参照加工 ＞ 道德判断加工									
1	1124	-1	45	7	5.92	-4	38	14	ACC
					5.68	-2	42	10	ACC
					3.67	10	52	4	R mid orbital gyrus
					3.60	-8	62	0	L mid orbital gyrus
					3.35	-6	64	2	L mid orbital gyrus
2	159	-27	51	23	3.46	-24	48	20	L middle frontal gyrus
					3.38	-28	48	22	L middle frontal gyrus
3	81	65	-40	20	2.38	66	-42	22	R superior temporal gyrus
4	34	-24	40	42	2.09	-26	42	40	L middle frontal gyrus
道德判断加工 ＞自我参照加工									
1	517	-49	-63	25	7.48	-48	-64	24	L Middle temporal gyrus
2	415	0	-57	32	7.82	0	-58	32	PCC
3	375	-5	52	34	4.63	-6	52	36	L Superior medial gyrus
					3.94	-2	50	34	
					2.71	0	62	26	
					2.16	6	60	20	
					1.72	4	62	16	
4	275	53	-57	27	3.53	58	-60	24	R Middle temporal gyrus
					3.38	54	-56	28	R Superior temporal gyrus
					3.35	56	-64	28	R Superior temporal gyrus
					3.26	52	-66	30	R Angular gyrus
5	80	-38	24	-23	2.24	-40	20	-26	L temporal pole
					1.85	-36	20	-22	L Temporal pole
自我参照加工 ∩ 道德判断加工									
1	75	1	55	-7	3.9	0	54	-6	L Mid orbital gyrus
2	54	-5	55	9	4.42	-6	56	8	L Superior medial gyrus

所有的解剖位置均通过 SPM Anatomy 工具进行定位.

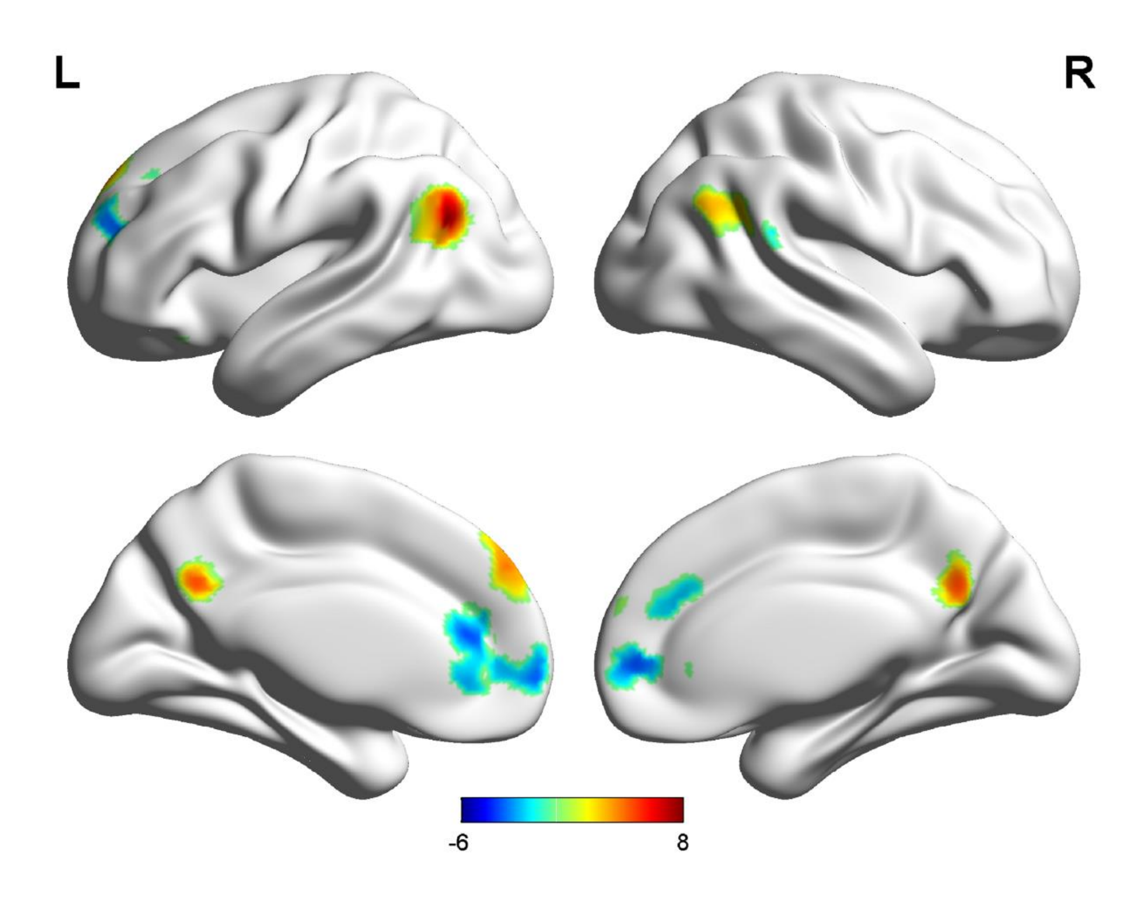


图 5.5　自我评估与道德评估元分析的对比分析。正值为道德 > 自我的脑区，负值为 自我 > 道德的脑区。

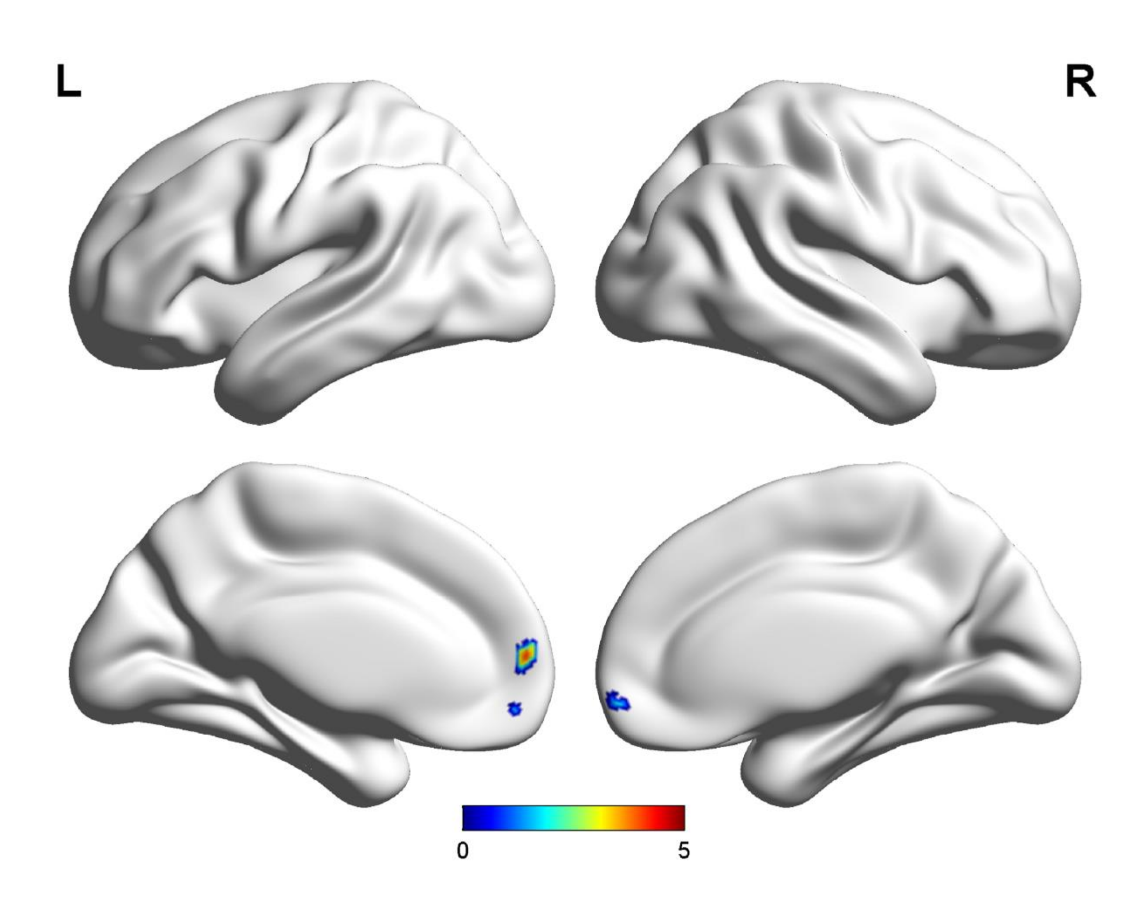


图 5.6 自我评估与道德评估神经成像共同激活的大脑区域。

5.3.4 讨论

本次元分析目的在于找到人们在进行自我评估时与进行道德判断在神经机制上的异同，从而探索道德自我的神经基础。通过使用基于坐标的元分析 ALE 方面，我们的结果显示，在自我评估的过程上，双侧的前扣带回/内侧前额叶（ACC/medial frontal gyrus）与左侧的额中回（middle frontal gyrus）一致地激活；道德判断任务则一致激活了双侧的内上额叶（superior medial frontal gyrus）、左侧 PCC、双侧的颞中回（middle temporal gyrus）、右侧颞上回（right superior temporal gyrus）、左侧眶额皮层（left mid orbital gyrus）。对于两个任务的比较分析发现，与道德判断任务相比，自我评估任务更地激活了双侧 ACC/mid orbital gyrus、左侧 middle frontal gyrus、右侧 superior temporal gyrus；而道德判断任务则更多地激活了而双侧的 middle temporal gyrus、PCC、左侧 temporal pole 和左侧的 superior medial frontal gyrus。

更重要的是，我们的联合分析表明，左侧背内侧前额叶和眶额皮层在两种条件下均激活，表明这两个大脑结构可能参与了自我加工与道德判断。这个结果与先前

自我增强效应的神经机制研究一致，表明眶额皮层以及内侧前额叶在自我概念的保持中有重要作用。例如，Hughes and Beer (2013)使用社会反馈的操纵，来告诉被试他们被同伴所接受或者不被同伴所接受，从而来威胁他们的自我概念，然后让他们对一组形容词进行自我参照任务（即判断该词是否适合来描述自己）。他们发现，当自我概念被威胁自己，被试的额极等区域更加激活，可能反映了对自我概念积极概念的保持。

本研究的结果发现，虽然道德判断与自我参照加工在任务上迥异，但是两者却共同激活了背内前额叶。这个结果可能表明，在道德判断过程和自我参照加工中，均有道德自我的内隐参照加工。但是，这一结论还需要更多的证据来支撑。

第六章　总讨论

6.1　主要研究结果总结

道德自我对于个体而言具有重要的意义，不仅与个体的道德行为有着紧密的联系，而且对个体的认知过程有着非常重要的影响。通过上述的 11 个研究，我们对道德自我在认知加工中的优势进行了系统地考察，同时也对其神经机制进行了初步的探索。7 个行为实验的结果一致表明，积极的道德自我信息能力促进知觉决策；两个神经成像的研究则表明，眶额叶皮层可能在道德自我中起到关键作用：

首先，积极的道德自我信息在社会联结任务中表现出了显示的优势。不管是在研究 4 中还是在研究 8 中，当道德自我与自我的其他方面以及他人进行比较时，均表现出明显的优势；

其次，通过行为数据的模型分析，表明积极道德自我信息的优势主要表现在信息的积累速率上，即道德自我信息的信息累积更快，这种信息累积的速率与物理上突显性比较相似，表明道德自我在知觉中的优势与物理上的突显性类似；

第三，使用不同的任务（研究 5、6）或者使用不同的分类标准（研究 9）均发现积极的道德自我信息具有优势，表明这种效应相对稳定；但是在外显任务中（研究 4、9）比在内隐任务（5、6）中的效应量更大；

第四，仅使用道德信息（研究 1~3）时与自我的不同方面（研究 4）表现出了一致的行为模式，说明当自我未作为外显的信息时，对信息的加工可能使用道德自我的内隐参照加工；

第五，研究 1~6 的元分析表明，道德信息对知觉的影响效应量稳定，在 *d* prime 表现为小的效应量，在反应时间上表现为中等的效应；

第六，使用自我报告的方法，发现被试在道德方面的自我增强效应显著高于其他领域（能力、社交能力和身体吸引力）；但 VBM 研究却未能发现道德自我评分与大脑灰质体积之间的相关。但是对先前神经成像的元分析却发现，前额叶尤其是眶额叶皮层，可能是道德自我功能相关的脑区。

6.2　道德自我的知觉优势

通过研究 1~3 表明，积极的道德信息在知觉加工中有促进作用，而研究 4~6 表明，这种积极道德信息的促进使用，本质上是由于道德自我的内隐参照加工。随后研究 8 和 9 的模型分析也一致地表明，道德自我在知觉过程中也存在着优势，这

种优势在外显任务情况下优势显著。这些结果首先验证了道德自我作为一种高级社会信息，确实能够对知觉过程产生影响。

其次，这些结果与先前关于道德自我记忆的结果一致，说明道德自我信息对于认知过程能够产生重要的影响，而且这种影响主要表明为对积极方面与消极方面的不对称。例如对道德事件的记忆效果的研究表明，人们倾向于记住自己做过的道德事件而容易忘记自己做过的不道德事件(Kouchaki & Gino, 2016; Sedikides et al., 2016)。但是，本研究将道德自我信息对认知加工的影响从较高级的记忆过程进一步扩散到知觉这个较为低级的认知过程，进一步表明了道德自我信息的重要作用。

从一个更加广义的角度来讲，本文的研究结果将社会心理学与认知心理学中关于自我的研究连接起来，证实了自我增强效应跨任务的一致性。道德是社会知觉的基本维度(Goodwin et al., 2014)，而道德自我是个体身份与认同的核心部分(Molouki & Bartels, 2017; Newman, Bloom, et al., 2014; Strohminger & Nichols, 2014)。正是如此，道德自我在自我概念中占据着核心的地位，而自我增强效应在道德自我概念方面表现最为明显。本研究重复了先前研究的结果，表明道德自我增强效应高于其他领域(Tappin & McKay, 2016)。另一方面，自我增强效应在认知过程中的表现已经在认知心理学有着非常广泛的研究，无论是在意识(Di et al., 2007; Macrae et al., 2017)、知觉(Ma & Han, 2010; Sui & Humphreys, 2013; Sui, Liu, Mevorach, & Humphreys, 2015)还是记忆(Rogers et al., 1977)，均发现对自相关信息的优先加工（self-prioritization）。本研究将社会心理学中道德自我概念的优势作用扩散到认知加工领域，发现道德自我信息在知觉中的加工也具有优势作用，而且这种优势作用表现为知觉上的突显性，与物理上的突显性具有相似的特点（研究 8 和 9）。

6.3　道德自我：促进还阻碍作用？

本文对道德自我的研究表明，道德自我在知觉中有明显的加工优势而对于非道德自我（中性的自我与不道德的自我），则加工相对较差，甚至可能比他人条件的反应慢、准确率更低（研究 4）。这个结果的模式一方面重复了先前研究中自我信息在知觉加工优势(Sui et al., 2012)，同时也进一步拓展了先前的研究，将自我优势的研究进行一步精细化。

实际上，先前已经有研究表明，当自我信息与消极信息时，被试的认知加工会更差。如前所述，Kouchaki and Gino (2016)发现，个体对于自己做过的不道德行为记忆效果更差。在自我信息的知觉加工研究中，Sui, Ohrling, and Humphreys (2016)发现，仅呈现消极心境的音乐，可以让自我信息的加工优势消除。因此，在消极或者中性自我道德信息中，可能是由于消极的效价导致无法获得加工的优势。

积极道德自我与消极/中性道德自我在认知加工上的特点有两种可能的解释。一种可能是自我本身具有积极效价，而积极道德本身具有促进信息加工的特点；因此积极道德自我能够获得快速的加工，是由于积极效价的原因。这一点与最近情绪、奖赏等方面刺激的知觉研究结果一致(Pool, Brosch, Delplanque, & Sander, 2016)，虽然自我信息与奖赏信息在知觉中的作用可能不一致(Sui, Yankouskaya, & Humphreys, 2015)。

另一种可能的解释是中性/消极道德自我具有威胁作用，这种威胁性阻碍了个体对信息进行快速有效的加工。例如，在记忆方面的研究显示，记忆本身可能作为一种维持积极自我的策略(Sedikides & Green, 2009)，对于消极的刺激记忆效果更差。这种保护性的机制是否也出现在知觉水平，还有待于进一步研究的验证。对于这些假设的验证，将有助于进一步理解道德自我加工的机制问题。

此外，道德自我的效应与本文研究 1、2、3 中对道德品质信息在知觉中的优势也表现出了相同的模式，而在他人的不同道德条件下，并未表现出这个模式。这个结果可能说明，道德信息在知觉决策中的作用与道德自我的不同方面相似。这可能是由于在道德信息没有明显参照时，个体会自动地使用自我参照来对信息进行加工，从而产生了道德自我不同方面相同的结果模式。这种自我投射理论（Self-Project Theory）在社会心理学的其他领域已经得到了不少证据的支持(Krueger, 2007)，尤其是在对他人的心理状态和动机的推断中(Jenkins & Mitchell, 2010; Tamir & Mitchell, 2013)。

6.4　道德自我的认知计算层次

本文研究的结果表明了道德自我具有知觉上的加工优势，但是对于这一现象的机制仅进行了初步的探讨。根据 David Marr 的计算层次模型(Marr, 2010)，任务认知功能的计算模型可以分为三个层次。第一个层次，也就是最外显的层次，是计算理论，即计算的目标是什么，为什么这个计算是合理的以及计算策略的逻辑是什么；第二个层次是表征和算法层次，即如何从算法角度来实施计算理论，输入和输出的表征是什么，以及输入与输出之间转换的算法是什么；第三个层次是硬件层次，即表征与算法需要在什么样的物理基础上进行实现（表 6.1）。

从社会认知的角度来讲，第一个和第三个层次对应的分别是功能与神经机制；而第二个层次则是从神经机制到心理功能的认知机制。具体到道德自我的加工优势上，这三个层次则是：道德自我在信息加工中的优势是什么以及为什么会有这个优势；道德自我的加工优势是通过什么心理机制/神经机制实现的；神经系统中实现这种加工优势的基础是什么？（见表 6.1）

表 6.1　计算视角下道德自我的三个层次

	计算理论	表征与算法	硬件实施
计算视觉	计算的目标、合理性及计算策略的逻辑	计算理论如何实现？ 输入与输出的表征是什么？ 输入与输出之间的转换的算法是什么	表征与算法如何在物理层面实现
道德自我	道德自我的优势是什么？为什么？	道德自我的加工优势的心理机制与神经机制是什么？	道德自我优势神经基础是什么？

从计算视角来看，本文对第一个层次的问题进行了系统地考察（图 6.1），发现了道德自我在信息加工中具有优势，表现为积极的道德自我信息与中性和消极的道德自我信息相比有着较强的加工优势。与前人的研究一起，本文的研究结果表明，道德自我加工的优势可能是由于维持积极的自我形象，从而保持自我概念的稳定性，维持个体的身体和心理的健康(Mann, Hosman, Schaalma, & de Vries, 2004; Taylor & Brown, 1988)。

本文对第二个层次的问题进行了初步的探索（图 6.1）。通过对行为数据的模型分析，本文发现道德自我信息的加工并非主要由于对道德自我信息存在偏见（扩散模型中的初始偏见 z），而是由于信息加工的速率变化。可能是由于道德自我中，自我相关性带来的整合效应(Sui & Humphreys, 2015a, 2016)。这种整合的非对称性（即对道德自我的优先的、快速的整合与对非道德自我的非优先整合）可能与个体的信念更新有类似的机制(Gino, Norton, & Weber, 2016; Sharot & Garrett, 2016)。

对于道德自我的硬件层次的分析，本文的研究也进行了初步的探索（图 6.1）。VBM 的研究未发现与道德自我评估相关的脑结构，但是我们对道德评估以及自我评估两类神经成像研究的元分析表明，前额叶可能是道德自我的关键脑区，尤其是眶额叶皮层。这与对行为数据的模型分析结果中信息累积速率的优势具有一致性，因为先前的研究表明，信息累积速度可能主要与额-顶网络有关(Mulder, van Maanen,

& Forstmann, 2014)。

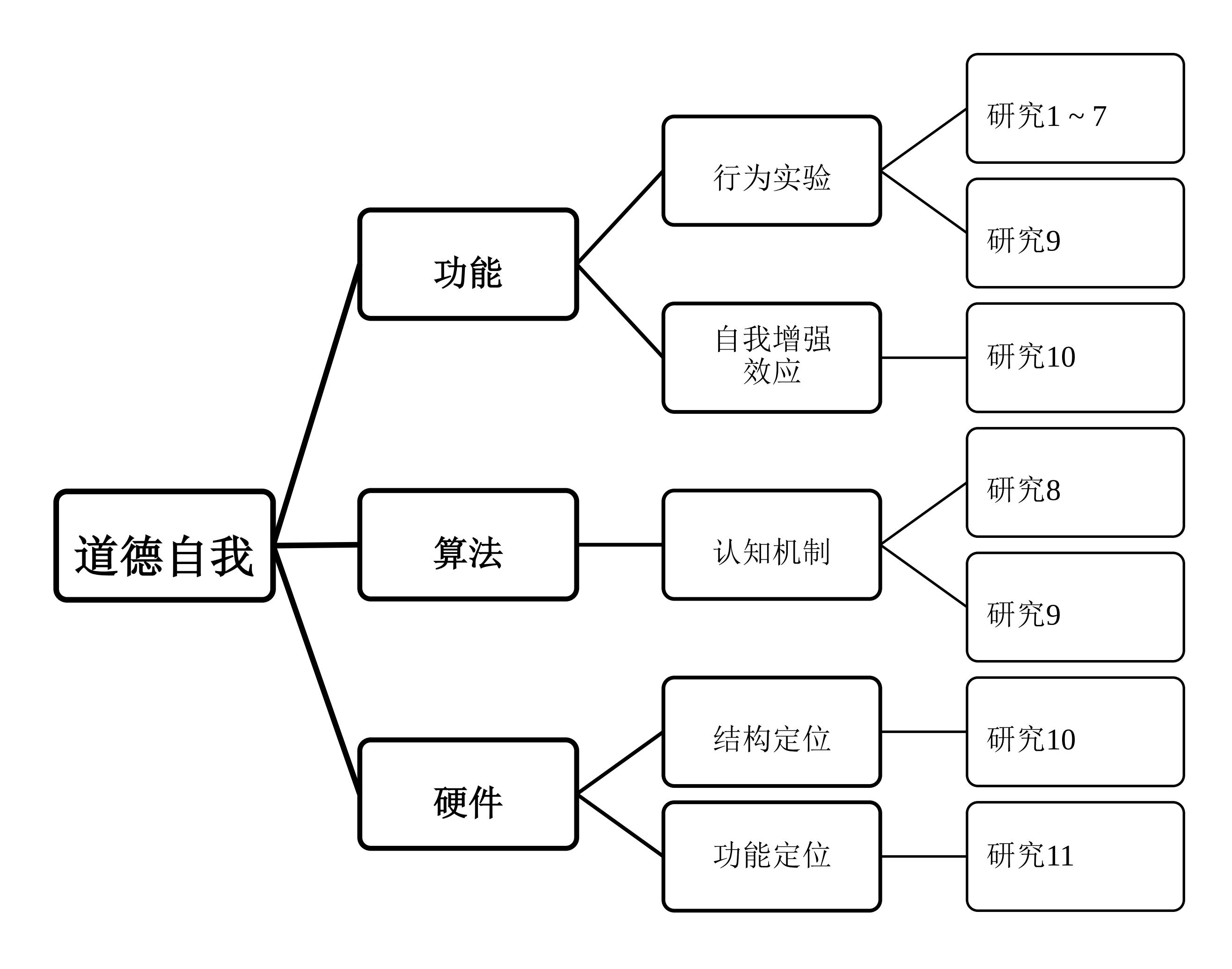


图 6.1　计算视角的道德自我研究

以上的分析表明，进一步深入研究道德自我的加工优势，需要进一步深入探讨其认知加工的机制以及认知计算模型。如图 6.2 所示，完整地理解功能、算法与硬件层次的关系，不仅需要进一步累积多方面的数据，还需要有良好的理论模型还整合三者之间的关系(Forstmann, Wagenmakers, Eichele, Brown, & Serences, 2011)。

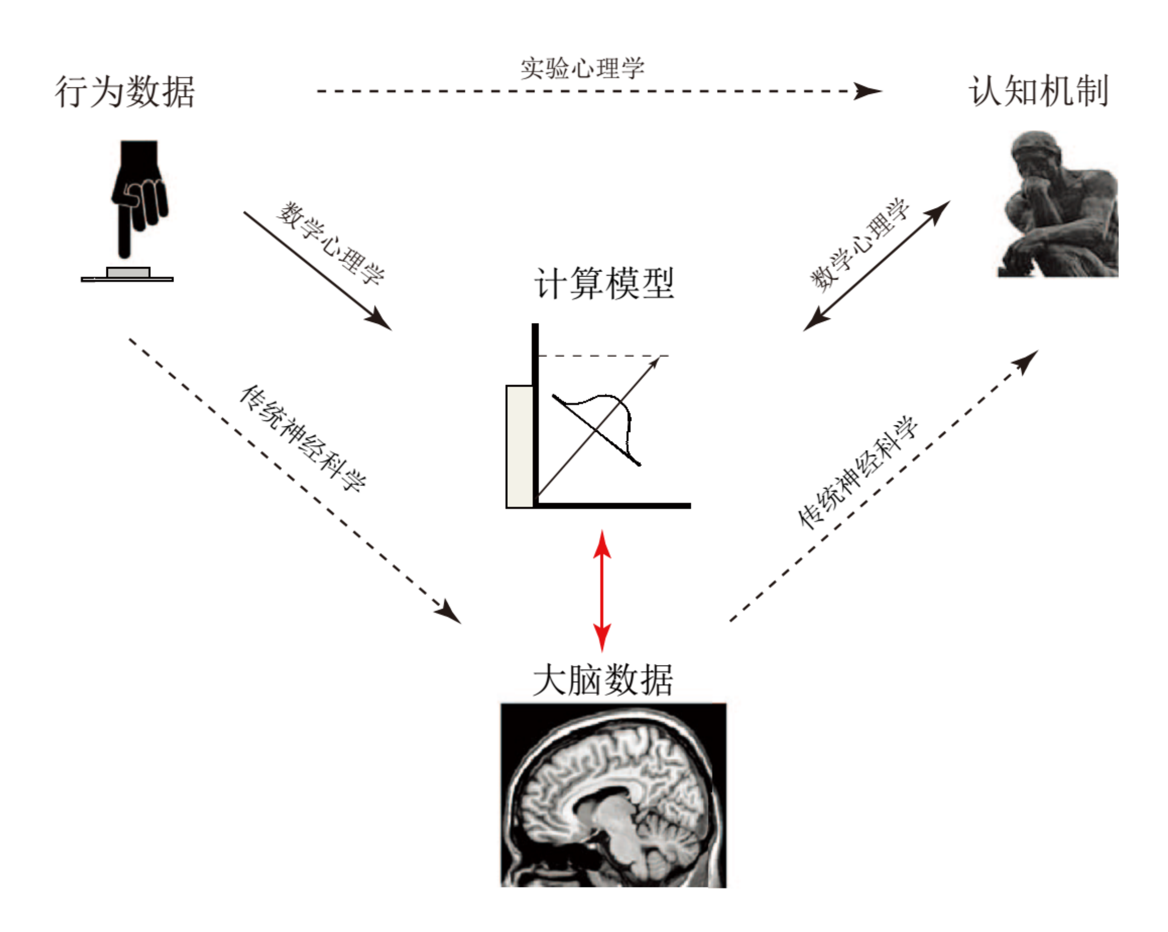


图 6.2 行为数据、认知过程与大脑数据之间的关系，需要通过正式的模型来进行联合。本图片翻译自(Forstmann et al., 2011)

6.5　研究的创新之处与局限

6.5.1　创新之处

道德自我对人类行为与道德判断均有着重要的影响。然而先前的研究一般是从行为层面进行的，包括自我报告或者行为任务等。然而，很少有研究深入探讨了其心理机制以及神经机制。本文从采用严格的实验室实验，对道德自我在知觉上的加工进行了系统地探索，从确认道德信息对知觉的影响（研究 1、2、3）到验证道德自我对知觉的外显和内隐影响（研究 4、5、6），再使用行为数据建模的方法，分析道德自我对知觉决策影响的机制，最后使用 VBM 和神经成像元分析的方法对道德自我进行探索。

本课题首次从认知计算的角度，对社会心理中的道德自我进行研究，将社会心理学与认知神经科学结合起来。使用计算模型的方法来研究社会认知神经科学是近年来社会认知神经科学的一个趋势(Stanley & Adolphs, 2013)。本文从计算模型的经典理论计算层次出发，从三个水平对道德自我进行了探讨，这对于今后继续深入研究社会认知背后的计算模型提供了基础。

本课题的研究结果扩展了关于自我加工优势的理论(Sui & Humphreys, 2015a, 2016)。正如身体自我与心理自我在心理与神经层次上有着区别(Hu et al., 2016)，在个体对自我的评估中，不同道德效价的自我，在认知加工中也不是使用相同策略进行加工的。

此外，本课题中对道德自我的神经机制的探索，一方面验证了先前关于自我增强效应的神经机制的研究(Hughes & Beer, 2013)；另一方面也为今后研究进行一步研究道德自我的神经机制提供了基础。

6.5.2 局限性与未来研究

本课题使用严格的实验室方法研究道德自我，能够良好地进行控制实验中的混淆变量，但是也不可避免地带来了一个问题，即将道德自我这个社会概念与其社会背景相剥离。被试仅在实验室中对着屏幕进行按键反应，而不是在真正的社会情境中进行道德决策，带来了明显的生态学效度的问题。因此，今后的研究需要加强道德自我研究的生态学效度。例如，可以采用博弈的方法，一方面仍然可以继续进行实验的控制，另一方面，增加了互动性。这将对于研究道德自我来说是一个很有潜力的取向。

其次，本研究的结果虽然试图从计算模型的三个层次来研究道德自我，但是在每个层次上的研究却未能深入。在功能层次上，我们发现了道德自我的加工优势，但是对于这种优势加工的原因却仍然未能给出确定的答案，在未来的研究中，我们可以探索如何能够调控道德自我的加工优势。在表征与算法层次，本课题发现了道德自我加工优势主要表现为信息累积速率的增加，但是具体使用何种算法能够实现这种信息累积速率的不对称，还有待进一步开发。最近，Gino et al. (2016)提出，可以采用动机性贝叶斯模型（motivated Bayesian model）来解释个体在道德行为与道德判断之间的不一致。道德行为包含了非常复杂的因素，实际上很难用精确的算法去模拟。但是在知觉层面，我们可以采用较为精细的算法去模拟这种道德自我在知觉决策中的优势，这也是未来一个非常值得探索的研究方法。

在神经基础层次，本研究仅利用结构数据和元分析数据进行了初步的探索，未能使用任务态的脑电实验、fMRI 实验对道德自我进行功能定位。更进一步，使用

采用多种方法研究道德自我的神经网络也将是今后需要进一步深入之处。总之，本课题只是进行了初步的探索，在三个层次上均有许多工作需要继续展开（见图 6.3）。

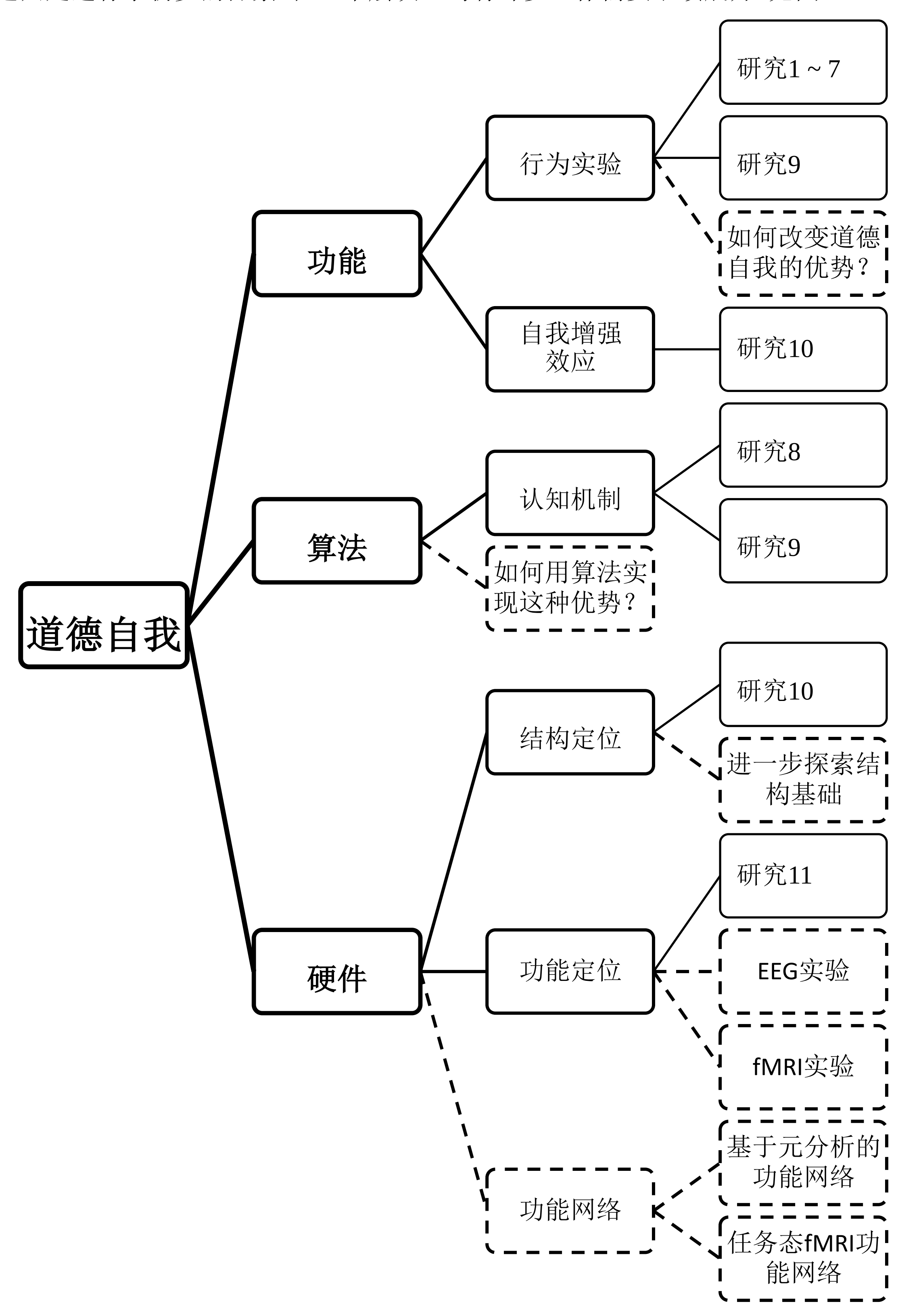


图 6.3　计算视角的道德自我的未来研究，虚线部分为将来可能进行的研究。

此外，本课题未能充分将计算模型的优势与神经成像的方法进行紧密的结果。例如，使用扩散模型可以根据数据结果对模型参数进行试次水平的估计，从而能够将每个模型参数在单个试次上的变化与神经信号的变化结合起来，更加细致地研究某个模型参数与神经机制之间的关系(Frank et al., 2015; Turner, Rodriguez, Norcia, McClure, & Steyvers, 2016)。采用这种研究方法，可以将认知机制与神经基础结合起来，进一步深入探索道德自我信息的加工中，神经活动与认知机制之间的关系。

对道德自我的研究，最终的目的在于理解人类道德行为背后的机制，从而为解决人类社会中存在的道德问题提供理论基础。道德自我稳定的加工优势，以及其在各个任务中的内隐参照作用，表明道德自我可能是人类许多道德判断的参照。当参照本身是一种虚假的高标准时（人们的道德自我优越感），人们容易产生不理性的对他人的道德批判。了解人类道德自我的加工机制，将可能为如何去改变这种默认的、自动的道德自我优越感而提供前期的研究基础。

6.6　结论

本课题结合行为实验、元分析、计算模型和神经成像的方法，系统地探索了道德自我在认知加工中的优势。这些研究能够有助于解决先前研究中关于道德信息在知觉加工中是否具有优势的争论(Firestone & Scholl, 2015; Gantman & Van Bavel, 2015, 2016)。更重要的是，本研究发现，积极的道德自我作为自我概念中核心的部分，在知觉中相对于道德他人或者中性/消极道德自我而言具有优势。具体而言，从本课题的 11 个研究中，可以得到如下结论：

第一、道德信息能够调节信息在知觉匹配中的加工，积极的道德品质信息具有加工的优势；

第二、与中性或者消极的自我道德信息以及他人信息相比，积极的道德自我信息在知觉加工中具有稳定的优势，这种优势表现在 *d* prime 上为中等偏小的效应，在反应时间上表现为中等偏大的效应量；

第三、道德自我信息在知觉中的加工优势具有跨任务的稳定性，可能反映了道德自我在信息加工中的锚定作用；

第四、道德自我信息在知觉加工中的优势主要是由于扩散速率上优势，这种优势具有跨任务的稳定性；而在非反应时间和初始偏见上的优势受到任务的调节；

第五、元分析结果表明，道德自我信息的加工可能在背内侧前额叶。

参考文献

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nat Rev Neurosci, 4*(3), 165-178. doi:10.1038/nrn1056

Adolphs, R. (2009). The Social Brain: Neural Basis of Social Knowledge. *The Annual Review of Psychology, 60*, 693-716. doi:10.1146/annurev.psych.60.110707.163514

Adolphs, R., Nummenmaa, L., Todorov, A., & Haxby, J. V. (2016). Data-driven approaches in the investigation of social perception. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 371*(1693). doi:10.1098/rstb.2015.0367

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. A. Dunning, & J. Krueger (Eds.), *The self in social judgment* (Vol. 1, pp. 85-106). New York: Psychology Press.

Alicke, M. D., Vredenburg, D. S., Hiatt, M., & Govorun, O. (2001). The "Better Than Myself Effect". *Motivation and Emotion, 25*(1), 7-22. doi:10.1023/a:1010655705069

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences, 4*(7), 267-278. doi:10.1016/S1364-6613(00)01501-1

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin, 111*(2), 256-274. doi:10.1037/0033-2909.111.2.256

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences, 108*(25), 10367-10371. doi:10.1073/pnas.1104047108

Anderson, E., Siegel, E. H., Bliss-Moreau, E., & Barrett, L. F. (2011). The Visual Impact of Gossip. *Science, 332*(6036), 1446-1448. doi:10.1126/science.1201574

Aquino, K., & Reed II, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology, 83*(6), 1423-1440. doi:10.1037/0022-3514.83.6.1423

Araujo, H. F., Kaplan, J., & Damasio, A. (2013). Cortical midline structures and autobiographical-self processes: an activation-likelihood estimation (ALE) meta-analysis. *Frontiers in Human Neuroscience, 7*. doi:10.3389/fnhum.2013.00548

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *NeuroImage, 38*(1), 95-113. doi:10.1016/j.neuroimage.2007.07.007

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods, 37*(3), 379-384. doi:10.3758/BF03192707

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General, 128*(1), 32-55. doi:10.1037/0096-3445.128.1.32

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language, 59*(4), 495-523. doi:10.1016/j.jml.2007.10.004

Bargh, J. A., McKenna, K. Y. A., & Fitzsimons, G. M. (2002). Can You See the Real Me? Activation and Expression of the "True Self" on the Internet. *Journal of Social Issues, 58*(1), 33-48. doi:10.1111/1540-4560.00247

Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, A. D. (1997). In a very

different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology, 72*(6), 1335-1348. doi:10.1037/0022-3514.72.6.1335

Batson, C. D., Thompson, E. R., Seuferling, G., Whitney, H., & Strongman, J. A. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology, 77*(3), 525-537. doi:10.1037/0022-3514.77.3.525

Baumeister, R. F., Zhang, L., & Vohs, K. D. (2004). Gossip as cultural learning. *Review of General Psychology, 8*(2), 111-121. doi:10.1037/1089-2680.8.2.111

Bennett, C. M., Miller, M. B., & Wolford, G. L. (2009). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: an argument for multiple comparisons correction. *NeuroImage, 47*(Supplement 1), S125. doi:10.1016/S1053-8119(09)71202-9

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn Reson Med, 34*(4), 537-541. doi:10.1002/mrm.1910340409

Blasi, A. (1983). Moral cognition and moral action: A theoretical perspective. *Developmental Review, 3*(2), 178-210. doi:10.1016/0273-2297(83)90029-1

Blasi, A. (1984). Moral Identity: Its Role in Moral Functioning. In W. M. Kurtines & J. J. Gewirtz (Eds.), *Morality, Moral Behavior and Moral Development* (pp. 128–139). New York: John Wiley & Sons.

Blasi, A. (2004). Moral functioning: Moral understanding and personality. In D. K. Lapsley & D. Narvaez (Eds.), *Moral Development, Self, and Idenity*. Mahwah, NJ: Lawrance Erlbaum & Associates.

Bostyn, D. H., & Roets, A. (2017). Trust, Trolleys and Social Dilemmas: A Replication Study. *Journal of Experimental Psychology: General, 146*(5), e1-e7. doi:10.1037/xge0000295

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433-436.

Brambilla, M., Hewstone, M., & Colucci, F. P. (2013). Enhancing moral virtues: Increased perceived outgroup morality as a mediator of intergroup contact effects. *Group Processes & Intergroup Relations, 16*(5), 648-657. doi:10.1177/1368430212471737

Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., & Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *British Journal of Social Psychology, 51*(1), 149-166. doi:10.1111/j.2044-8309.2010.02011.x

Brett, M., Anton, J.-L., Valabregue, R., & Poline, J.-B. (2002, June 2-6). *Region of interest analysis using the MarsBar toolbox for SPM 99.* Paper presented at the the 8th International Conference on Functional Mapping of the Human Brain, Sendai, Japan.

Brown, J. D. (2012). Understanding the Better Than Average Effect. *Personality and Social Psychology Bulletin, 38*(2), 209-219. doi:doi:10.1177/0146167211432763

Carazza, B. (1977). The history of the random-walk problem: considerations on the interdisciplinarity in modern physics. *La Rivista del Nuovo Cimento (1971-1977), 7*(3), 419-427. doi:10.1007/bf02747280

Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General, 143*(4), 1476-1488. doi:10.1037/a0035813

Chiong, W., Wilson, S. M., D'Esposito, M., Kayser, A. S., Grossman, S. N., Poorzand, P., . . . Rankin, K. P. (2013). The salience network causally influences default mode network activity during moral reasoning. *Brain, 136*(6), 1929-1941. doi:10.1093/brain/awt066

Cole, Michael W., Bassett, Danielle S., Power, Jonathan D., Braver, Todd S., & Petersen, Steven E. (2014). Intrinsic and Task-Evoked Network Architectures of the Human Brain. *Neuron, 83*(1), 238-251.

doi:10.1016/j.neuron.2014.05.014
Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2 ed.). New York: Sage.
Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies with the Wason selection task. *Cognition, 31*(3), 187-276. doi:10.1016/0010-0277(89)90023-1
Cosmides, L., & Tooby, J. (2000). The cognitive neuroscience of social reasoning. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2 ed., pp. 1259-1272). Cambridge, MA: MIT Press.
Cousineau, D., Brown, S., & Heathcote, A. (2004). Fitting distributions using maximum likelihood: Methods and packages. *Behavior Research Methods, Instruments, & Computers, 36*(4), 742-756. doi:10.3758/bf03206555
Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review, 108*(3), 593-623. doi:10.1037/0033-295X.108.3.593
Cushman, F., Young, L., & Hauser, M. (2006). The Role of Conscious Reasoning and Intuition in Moral Judgment. *Psychological Science, 17*(12), 1082-1089. doi:10.1111/j.1467-9280.2006.01834.x
de Gelder, B., van Honk, J., & Tamietto, M. (2011). Emotion in the brain: of low roads, high roads and roads less travelled. *Nat Rev Neurosci, 12*(7), 425-425. doi:10.1038/nrn2920-c1
De Waal, F. (2005). *Our inner ape: The best and worse of human nature*. New York: Riverhead Books.
Di, H. B., Yu, S. M., Weng, X. C., Laureys, S., Yu, D., Li, J. Q., . . . Chen, Y. Z. (2007). Cerebral response to patient's own name in the vegetative and minimally conscious states. *Neurology, 68*(12), 895-899.
Dunbar, R. I. M. (2004). Gossip in evolutionary perspective. *Review of General Psychology, 8*(2), 100-110. doi:10.1037/1089-2680.8.2.100
Eastwood, J. D., Smilek, D., & Merikle, P. M. (2003). Negative facial expression captures attention and disrupts performance. *Perception & Psychophysics, 65*(3), 352-358.
Effron, D. A., & Conway, P. (2015). When virtue leads to villainy: advances in research on moral self-licensing. *Current Opinion in Psychology, 6*, 32-35. doi:10.1016/j.copsyc.2015.03.017
Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage, 59*(3), 2349-2361. doi:10.1016/j.neuroimage.2011.09.017
Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2006). Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. *NeuroImage, 32*(2), 570-582. doi:10.1016/j.neuroimage.2006.04.204
Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Hum Brain Mapp, 30*(9), 2907-2926. doi:10.1002/hbm.20718
Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M.-H., Evans, A. C., Zilles, K., & Amunts, K. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *NeuroImage, 36*(3), 511-521. doi:10.1016/j.neuroimage.2007.03.060
Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., & Zilles, K. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage, 25*(4), 1325-1335. doi:10.1016/j.neuroimage.2004.12.034
Epley, N., & Dunning, D. (2000). Feeling "holier than thou": Are self-serving assessments produced by errors in self- or social prediction? *Journal of Personality and Social Psychology, 79*(6), 861-875.

doi:10.1037/0022-3514.79.6.861

Everett, J. A. C., Pizarro, D. A., & Crockett, M. J. (2016). Inference of Trustworthiness From Intuitive Moral Judgments. *Journal of Experimental Psychology: General, 145*(6), 772-787. doi:10.1037/xge0000165

Feller, W. (1968). *An introduction to probability theory and its applications: volume I* (3 ed. Vol. 1). New York: John Wiley & Sons.

Feller, W. (1971). *An introduction to probability theory and its applications: volume 2* (Vol. 2). New York, U.S.: John Wiley & Sons.

Firestone, C., & Scholl, B. J. (2014). Enhanced visual awareness for morality and pajamas? Perception vs. memory in 'top-down' effects. *Cognition*. doi:10.1016/j.cognition.2014.10.014

Firestone, C., & Scholl, B. J. (2015). 'Moral Perception' Reflects Neither Morality Nor Perception. *Trends in Cognitive Sciences, 20*(2), 75-76. doi:10.1016/j.tics.2015.10.006

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83. doi:10.1016/j.tics.2006.11.005

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential Sampling Models in Cognitive Neuroscience: Advantages, Applications, and Extensions. *Annual Review of Psychology, 67*(1), 641-666. doi:10.1146/annurev-psych-122414-033645

Forstmann, B. U., Wagenmakers, E.-J., Eichele, T., Brown, S., & Serences, J. T. (2011). Reciprocal relations between cognitive neuroscience and formal cognitive models: opposites attract? *Trends in Cognitive Sciences, 15*(6), 272-279. doi:10.1016/j.tics.2011.04.002

Fouragnan, E., Chierchia, G., Greiner, S., Neveu, R., Avesani, P., & Coricelli, G. (2013). Reputational Priors Magnify Striatal Responses to Violations of Trust. *The Journal of Neuroscience, 33*(8), 3602-3611. doi:10.1523/jneurosci.3086-12.2013

Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., & Badre, D. (2015). fMRI and EEG Predictors of Dynamic Decision Parameters during Human Reinforcement Learning. *The Journal of Neuroscience, 35*(2), 485-494. doi:10.1523/jneurosci.2036-14.2015

Gantman, A. P., & Bavel, J. J. V. (2016). Exposure to Justice Diminishes Moral Perception. *Journal of Experimental Psychology: General, 145*(12), 1728-1739. doi:10.1037/xge0000241

Gantman, A. P., & Van Bavel, J. J. (2014). The moral pop-out effect: Enhanced perceptual awareness of morally relevant stimuli. *Cognition, 132*(1), 22-29. doi:10.1016/j.cognition.2014.02.007

Gantman, A. P., & Van Bavel, J. J. (2015). Moral Perception. *Trends in Cognitive Sciences, 19*(11), 631-633. doi:10.1016/j.tics.2015.08.004

Gantman, A. P., & Van Bavel, J. J. (2016). See for Yourself: Perception Is Attuned to Morality. *Trends in Cognitive Sciences, 20*(2), 76-77. doi:10.1016/j.tics.2015.12.001

Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science, 7*(4), 457-472.

Gilligan, C. (1982). *In a different voice*. Cambridge, Massachusetts: Harvard University Press.

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and Differentiation in Unethical Behavior. *Psychological Science, 20*(3), 393-398. doi:10.1111/j.1467-9280.2009.02306.x

Gino, F., Norton, M. I., & Weber, R. A. (2016). Motivated Bayesians: Feeling Moral While Acting Egoistically. *Journal of Economic Perspectives, 30*(3), 189-212. doi:10.1257/jep.30.3.189

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini Meta-Analysis of Your Own Studies: Some Arguments on Why and a Primer on How. *Social and Personality Psychology Compass, 10*(10), 535-549.

doi:10.1111/spc3.12267

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology, 106*(1), 148-168. doi:10.1037/a0034726

Goulden, N., Khusnulina, A., Davis, N. J., Bracewell, R. M., Bokde, A. L., McNulty, J. P., & Mullins, P. G. (2014). The salience network is responsible for switching between the default mode network and the central executive network: Replication from DCM. *NeuroImage, 99*, 180-190. doi:10.1016/j.neuroimage.2014.05.052

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & Ditto, P. H. (2012). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology, Forthcoming, 47*, 55-130.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and Conservatives Rely on Different Sets of Moral Foundations. *Journal of Personality and Social Psychology,, 96*(5), 1029-1046. doi:10.1037/a0015141

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*(2), 366-385. doi:10.1037/a0021847

Greene, J. D. (2014). The cognitive neuroscience of moral judgment and decision making. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences* (V ed.). Cambridge: MIT Press.

Greene, J. D. (2015). The rise of moral cognition. *Cognition*. doi:10.1016/j.cognition.2014.11.018

Greene, J. D., & Paxton, J. M. (2009). Patterns of neural activity associated with honest and dishonest moral decisions. *Proceedings of the National Academy of Sciences, 106*(30), 12506-12511. doi:10.1073/pnas.0900152106

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science, 293*(5537), 2105-2108. doi:10.1126/science.1062872

Greicius, M. D., Krasnow, B., Reiss, A. L., & Menon, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proceedings of the National Academy of Sciences, 100*(1), 253-258. doi:10.1073/pnas.0135058100

Haidt, J. (2001). The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment. *Psychological Review, 108*(4), 814-834. doi:10.1037/0033-295X.108.4.814

Haidt, J. (2007). The New Synthesis in moral Psychology. *Science, 316*(5827), 998-1002. doi:10.1126/science.1137651

Haidt, J., & Kesebir, S. (2010). Morality. In S. T. Fiske & D. Gilbert (Eds.), *The Handbook of Social Psychology* (5 ed.).

Harrison, B. J., Pujol, J., López-Solà, M., Hernández-Ribas, R., Deus, J., Ortiz, H., . . . Cardoner, N. (2008). Consistency and functional specialization in the default mode brain network. *Proceedings of the National Academy of Sciences, 105*(28), 9781-9786. doi:10.1073/pnas.0711791105

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2002). Quantile maximum likelihood estimation of response time distributions. *Psychonomic Bulletin & Review, 9*(2), 394-401. doi:10.3758/bf03196299

Heathcote, A., Popiel, S. J., & Mewhort, D. J. (1991). Analysis of response time distributions: An example using the Stroop task. *Psychological Bulletin, 109*(2), 340-347. doi:10.1037/0033-2909.109.2.340

Hedger, N., Gray, K. L. H., Garner, M., & Adams, W. J. (2016). Are visual threats prioritized without awareness? A critical review and meta-analysis involving 3 behavioral paradigms and 2696 observers. *Psychological Bulletin, 142*(9), 934-968. doi:10.1037/bul0000054

Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review, 94*(3), 319-340. doi:10.1037/0033-295X.94.3.319

Hofmann, W., Wisneski, D. C., Brandt, M. J., & Skitka, L. J. (2014). Morality in everyday life. *Science, 345*(6202), 1340-1343. doi:10.1126/science.1251560

Hohle, R. H. (1965). Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology, 69*(4), 382-386. doi:10.1037/h0021740

Hu, C., Di, X., Eickhoff, S. B., Zhang, M., Peng, K., Guo, H., & Sui, J. (2016). Distinct and common aspects of physical and psychological self-representation in the brain: A meta-analysis of self-bias in facial and self-referential judgements. *Neurosci Biobehav Rev*, *61*, 197–207. doi:10.1016/j.neubiorev.2015.12.003

Hughes, B. L., & Beer, J. S. (2013). Protecting the Self: The Effect of Social-evaluative Threat on Neural Representations of Self. *Journal of Cognitive Neuroscience, 25*(4), 613-622. doi:10.1162/jocn_a_00343

James, W. (1890). *The principles of psychology*. Henry Holt: New York.

Jenkins, A. C., & Mitchell, J. P. (2010). Mentalizing under Uncertainty: Dissociated Neural Responses to Ambiguous and Unambiguous Mental State Inferences. *Cerebral Cortex*, *20*(2), 404-410. doi:10.1093/cercor/bhp109

Jiang, Y., Costello, P., Fang, F., Huang, M., & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences, 103*(45), 17048-17052. doi:10.1073/pnas.0605678103

Jiang, Y., Costello, P., & He, S. (2007). Processing of Invisible Stimuli: Advantage of Upright Faces and Recognizable Words in Overcoming Interocular Suppression. *Psychological Science, 18*(4), 349-355. doi:10.1111/j.1467-9280.2007.01902.x

Jordan, A. H., & Monin, B. (2008). From Sucker to Saint. *Psychological Science*, *19*(8), 809-815. doi:10.1111/j.1467-9280.2008.02161.x

Jordan, J., Leliveld, M. C., & Tenbrunsel, A. (2015). The Moral Self-Image Scale: Measuring and Understanding the Malleability of the Moral Self. *Frontiers in Psychology*, *6*. doi:10.3389/fpsyg.2015.01878

Jordan, J. J., Hoffman, M., Nowak, M. A., & Rand, D. G. (2016). Uncalculating cooperation is used to signal trustworthiness. *Proceedings of the National Academy of Sciences, 113*(31), 8658–8663. doi:10.1073/pnas.1601280113

Kennedy, D., & Norman, C. (2005). What Don't We Know? *Science*, *309*(5731), 75-75. doi:10.1126/science.309.5731.75

Kirby, K., & Gerlanc, D. (2013). BootES: An R package for bootstrap confidence intervals on effect sizes. *Behavior Research Methods, 45*(4), 905-927. doi:10.3758/s13428-013-0330-5

Kitayama, S., King, A., Yoon, C., Tompson, S., Huff, S., & Liberzon, I. (2014). The Dopamine D4 Receptor Gene (DRD4) Moderates Cultural Difference in Independent Versus Interdependent Social Orientation. *Psychological Science*. doi:10.1177/0956797614528338

Klein, N., & Epley, N. (2016). Maybe holier, but definitely less evil, than you: Bounded self-righteousness in social judgment. *Journal of Personality and Social Psychology, 110*(5), 660-674. doi:10.1037/pspa0000050

Knobe, J. (2010). Person as scientist, person as moralist. *BEHAVIORAL AND BRAIN SCIENCES, 33*(04), 315-329. doi:10.1017/S0140525X10000907

Knutson, K. M., Krueger, F., Koenigs, M., Hawley, A., Escobedo, J. R., Vasudeva, V., . . . Grafman, J. (2010). Behavioral norms for condensed moral vignettes. *Social Cognitive and Affective Neuroscience, 5*(4), 378-384. doi:10.1093/scan/nsq005

Kohlberg, L. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347-480). Chicago, IL: Rand McNally.

Kouchaki, M., & Gino, F. (2016). Memories of unethical actions become obfuscated over time. *Proceedings of the National Academy of Sciences, 113*(22), 6166-6171. doi:10.1073/pnas.1523586113

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron, 93*(3), 480-490. doi:10.1016/j.neuron.2016.12.041

Krueger, J. I. (2007). From social projection to social behaviour. *European Review of Social Psychology, 18*(1), 1-35. doi:10.1080/10463280701284645

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863. doi:10.3389/fpsyg.2013.00863

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. Oxford, England: Academic Press.

Lancaster, J. L., Tordesillas-Gutiérrez, D., Martinez, M., Salinas, F., Evans, A., Zilles, K., . . . Fox, P. T. (2007). Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. *Hum Brain Mapp, 28*(11), 1194-1205. doi:10.1002/hbm.20345

Landau, M. J., Vess, M., Arndt, J., Rothschild, Z. K., Sullivan, D., & Atchley, R. A. (2011). Embodied metaphor and the "true" self: Priming entity expansion and protection influences intrinsic self-expressions in self-perceptions and interpersonal behavior. *Journal of Experimental Social Psychology, 47*(1), 79-87. doi:10.1016/j.jesp.2010.08.012

Lawrence, M. A. (2015). Ez: Easy analysis and visualization of factorial experiments. Retrieved from https://CRAN.R-project.org/package=ez

Leshikar, E., & Duarte, A. (2014). Medial prefrontal cortex supports source memory for self-referenced materials in young and older adults. *Cognitive, Affective, & Behavioral Neuroscience, 14*(1), 236-252. doi:10.3758/s13415-013-0198-y

Lieberman, M. D. (2007). Social Cognitive Neuroscience: A Review of Core Processes. *Annual Review of Psychology, 58*(1), 259-289. doi:10.1146/annurev.psych.58.110405.085654

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika, 40*(1), 77-105. doi:10.1007/bf02291481

Ma, Y., & Han, S. (2010). Why we respond faster to the self than to others? An implicit positive association theory of self-advantage during implicit face recognition. *Journal of Experimental Psychology: Human Perception and Performance, 36*(3), 619-633. doi:10.1037/a0015797

MacDonald, G., Saltzman, J. L., & Leary, M. R. (2003). Social approval and trait self-esteem. *Journal of Research in Personality, 37*(2), 23-40. doi:10.1016/S0092-6566(02)00531-7

Macrae, C. N., Visokomogilski, A., Golubickis, M., Cunningham, W. A., & Sahraie, A. (2017). Self-relevance prioritizes access to visual awareness. *Journal of Experimental Psychology: Human Perception and Performance, 43*(3), 438-443. doi:10.1037/xhp0000361

Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of*

the National Academy of Sciences, 97(8), 4398-4403. doi:10.1073/pnas.070039597

Mann, M., Hosman, C. M. H., Schaalma, H. P., & de Vries, N. K. (2004). Self-esteem in a broad-spectrum approach for mental health promotion. *Health Education Research, 19*(4), 357-372. doi:10.1093/her/cyg041

Marr, D. (2010). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: The MIT Press.

Massidda, D. (2013). Retimes: Reaction time analysis. Retrieved from https://cran.r-project.org/web/packages/retimes/index.html

Matzke, D., & Wagenmakers, E.-J. (2009). Psychological interpretation of the ex-Gaussian and shifted Wald parameters: A diffusion model analysis. *Psychonomic Bulletin & Review, 16*(5), 798-817. doi:10.3758/pbr.16.5.798

Mazar, N., Amir, O., & Ariely, D. (2008). The Dishonesty of Honest People: A Theory of Self-Concept Maintenance. *Journal of Marketing Research, 45*(6), 633-644. doi:10.1509/jmkr.45.6.633

McCarthy, R. J., & Skowronski, J. J. (2011). What will Phil do next?: Spontaneously inferred traits influence predictions of behavior. *Journal of Experimental Social Psychology, 47*(2), 321-332. doi:10.1016/j.jesp.2010.10.015

McCoy, B., & Theeuwes, J. (2016). Effects of reward on oculomotor control. *Journal of Neurophysiology, 116*(5), 2453-2466. doi:10.1152/jn.00498.2016

McDonald, A. R., Muraskin, J., Dam, N. T. V., Froehlich, C., Puccio, B., Pellman, J., . . . Craddock, R. C. (2017). The Real-time fMRI Neurofeedback Based Stratification of Default Network Regulation Neuroimaging Data Repository. *NeuroImage, 146*, 157-170. doi:10.1016/j.neuroimage.2016.10.048

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends in Cognitive Sciences, 15*(10), 483-506. doi:10.1016/j.tics.2011.08.003

Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature Reviews: Neuroscience, 6*(10), 799-809. doi:10.1038/nrn1768

Möller, J., & Savyon, K. (2003). Not Very Smart, thus Moral: Dimensional Comparisons Between Academic Self-Concept and Honesty. *Social Psychology of Education, 6*(2), 95-106. doi:10.1023/a:1023247910033

Molouki, S., & Bartels, D. M. (2017). Personal change and the continuity of the self. *Cognitive Psychology, 93*, 1-17. doi:10.1016/j.cogpsych.2016.11.006

Monin, B., & Jordan, A. H. (2009). The dynamic moral self: A social psychological perspective. In D. Narvaez & D. K. Lapsley (Eds.), *Personality, identity, and character: Explorations in moral psychology* (pp. 341-354). Cambridge, UK: University Press.

Monin, B., Pizarro, D. A., & Beer, J. S. (2007). Deciding Versus Reacting: Conceptions of Moral Judgment and the Reason-Affect Debate. *Review of General Psychology, 11*(2), 99-111. doi:10.1037/1089-2680.11.2.99

Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The Rejection of Moral Rebels: Resenting Those Who Do the Right Thing. *Journal of Personality and Social Psychology, 95*(1), 76-93. doi:10.1037/0022-3514.95.1.76

Moutsopoulou, K., & Waszak, F. (2012). Across-task priming revisited: Response and task conflicts disentangled using ex-Gaussian distribution analysis. *Journal of Experimental Psychology: Human Perception and Performance, 38*(2), 367-374. doi:10.1037/a0025858

Mulder, M. J., van Maanen, L., & Forstmann, B. U. (2014). Perceptual Decision Neuroscience - A Model-

Based Review. *Neuroscience, 277*, 872-884. doi:10.1016/j.neuroscience.2014.07.031

Mullen, E., & Monin, B. (2016). Consistency Versus Licensing Effects of Past Moral Behavior. *Annual Review of Psychology, 67*(1), 363-385. doi:10.1146/annurev-psych-010213-115120

Murray, R. J., Schaer, M., & Debbané, M. (2012). Degrees of separation: A quantitative neuroimaging meta-analysis investigating self-specificity and shared neural activation between self- and other-reflection. *Neuroscience & Biobehavioral Reviews, 36*(3), 1043-1059. doi:10.1016/j.neubiorev.2011.12.013

Nakane, T., Miyakoshi, M., Nakai, T., & Naganawa, S. (2016). How the Non-attending Brain Hears Its Owner's Name. *Cerebral Cortex, 26*(10), 3889-3904. doi:10.1093/cercor/bhv184

Narvaez, D., Lapsley, D. K., Hagele, S., & Lasky, B. (2006). Moral chronicity and social information processing: Tests of a social cognitive approach to the moral personality. *Journal of Research in Personality, 40*(6), 966-985. doi:10.1016/j.jrp.2005.10.003

Navarro, D. J., & Fuss, I. G. (2009). Fast and accurate calculations for first-passage times in Wiener diffusion models. *Journal of Mathematical Psychology, 53*(4), 222-230. doi:10.1016/j.jmp.2009.02.003

Newman, G. E., Bloom, P., & Knobe, J. (2014). Value Judgments and the True Self. *Personality and Social Psychology Bulletin, 40*(2), 203-216. doi:10.1177/0146167213508791

Newman, G. E., De Freitas, J., & Knobe, J. (2014). Beliefs About the True Self Explain Asymmetries Based on Moral Judgment. *Cognitive Science*, n/a-n/a. doi:10.1111/cogs.12134

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., & Panksepp, J. (2006). Self-referential processing in our brain—A meta-analysis of imaging studies on the self. *NeuroImage, 31*(1), 440-457. doi:10.1016/j.neuroimage.2005.12.002

Öhman, A., Flykt, A., & Esteves, F. (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General, 130*(3), 466-478. doi:10.1037/0096-3445.130.3.466

Park, H.-J., & Friston, K. (2013). Structural and Functional Brain Networks: From Connections to Cognition. *Science, 342*(6158). doi:10.1126/science.1238411

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437-442.

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafo, M. R., . . . Yarkoni, T. (2017). Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat Rev Neurosci*. doi:10.1038/nrn.2016.167

Pool, E., Brosch, T., Delplanque, S., & Sander, D. (2016). Attentional Bias for Positive Emotional Stimuli: A Meta-Analytic Investigation. *Psychological Bulletin, 142*(1), 79-106. doi:10.1037/bul0000026

Pronin, E. (2008). How We See Ourselves and How We See Others. *Science, 320*(5880), 1177-1180. doi:10.1126/science.1154199

Pruessmann, K. P., Weiger, M., Scheidegger, M. B., & Boesiger, P. (1999). SENSE: Sensitivity encoding for fast MRI. *Magnetic resonance in medicine, 42*(5), 952-962. doi:10.1002/(SICI)1522-2594(199911)42:5<952::AID-MRM16>3.0.CO;2-S

Qin, P., & Northoff, G. (2011). How is our self related to midline regions and the default-mode network? *NeuroImage, 57*(3), 1221-1233. doi:10.1016/j.neuroimage.2011.05.028

R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna , Austria. Retrieved from https://www.R-project.org/

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59-108. doi:10.1037/0033-

295X.85.2.59

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86*(3), 446-461. doi:10.1037/0033-2909.86.3.446

Ratcliff, R. (2002). A diffusion model account of response time and accuracy in a brightness discrimination task: Fitting real data and failing to fit fake but plausible data. *Psychonomic Bulletin & Review, 9*(2), 278-291. doi:10.3758/bf03196283

Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision, 2*(4), 237-279. doi:10.1037/dec0000030

Ratcliff, R., & McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Computation, 20*(4), 873-922. doi:10.1162/neco.2008.12-06-420

Ratcliff, R., & Rouder, J. N. (1998). Modeling Response Times for Two-Choice Decisions. *Psychological Science, 9*(5), 347-356. doi:10.1111/1467-9280.00067

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences, 20*(4), 260-281. doi:10.1016/j.tics.2016.01.007

Richard, F. D., Bond Jr, C. F., & Stokes-Zoota, J. J. (2003). One Hundred Years of Social Psychology Quantitatively Described. *Review of General Psychology, 7*(4), 331-363. doi:10.1037/1089-2680.7.4.331

Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology, 35*(9), 677-688. doi:10.1037/0022-3514.35.9.677

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

Rule, N. O., Ambady, N., & Adams Jr, R. B. (2009). Personality in perspective: Judgmental consistency across orientations of the face. *Perception, 38*, 1688–1699.

Schmiedek, F., Oberauer, K., Wilhelm, O., Süß, H.-M., & Wittmann, W. W. (2007). Individual differences in components of reaction time distributions and their relations to working memory and intelligence. *Journal of Experimental Psychology: General, 136*(3), 414-429. doi:10.1037/0096-3445.136.3.414

Sedikides, C., & Green, J. D. (2009). Memory as a Self-Protective Mechanism. *Social and Personality Psychology Compass, 3*(6), 1055-1068. doi:10.1111/j.1751-9004.2009.00220.x

Sedikides, C., Green, J. D., Saunders, J., Skowronski, J. J., & Zengel, B. (2016). Mnemic neglect: Selective amnesia of one’s faults. *European Review of Social Psychology, 27*(1), 1-62. doi:10.1080/10463283.2016.1183913

Seeley, W. W., Menon, V., Schatzberg, A. F., Keller, J., Glover, G. H., Kenna, H., . . . Greicius, M. D. (2007). Dissociable Intrinsic Connectivity Networks for Salience Processing and Executive Control. *The Journal of Neuroscience, 27*(9), 2349-2356. doi:10.1523/jneurosci.5587-06.2007

Sharot, T., & Garrett, N. (2016). Forming Beliefs: Why Valence Matters. *Trends in Cognitive Sciences, 20*(1), 25-33. doi:10.1016/j.tics.2015.11.002

Shore, D. M., & Heerey, E. A. (2013). Do social utility judgments influence attentional processing? *Cognition, 129*(1), 114-122. doi:10.1016/j.cognition.2013.06.011

Shweder, R. A., Much, N. C., Mahapatra, M., & Park, L. (1997). The “big three” of morality (autonomy, community, and divinity), and the “big three” explanations of suffering. In A. Brandt & P. Rozin (Eds.), *Morality and health* (pp. 119-169). New York, NY: Routledge.

Smith, P. L. (2000). Stochastic Dynamic Models of Response Time and Accuracy: A Foundational Primer. *Journal of Mathematical Psychology, 44*(3), 408-463. doi:10.1006/jmps.1999.1260

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583-639. doi:10.1111/1467-9868.00353

Stanley, D. A., & Adolphs, R. (2013). Toward a Neural Basis for Social Behavior. *Neuron, 80*(3), 816-826. doi:10.1016/j.neuron.2013.10.038

Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition, 114*(3), 447-454. doi:10.1016/j.cognition.2009.11.003

Sternberg, S. (1964, October). *Estimating the distribution of additive reaction-time components.* Paper presented at the the meeting of the Psychonomic Society, Niagara Falls, Ontario, Canada.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist, 57*(4), 421-457.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika, 25*(3), 251-260. doi:10.1007/bf02289729

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition, 131*(1), 159-171. doi:10.1016/j.cognition.2013.12.005

Sui, J., He, X., & Humphreys, G. W. (2012). Perceptual effects of social salience: Evidence from self-prioritization effects on perceptual matching. *Journal of Experimental Psychology: Human Perception and Performance, 38*(5), 1105-1117. doi:10.1037/a0029792

Sui, J., & Humphreys, G. W. (2013). The boundaries of self face perception: Response time distributions, perceptual categories, and decision weighting. *Visual Cognition, 21*(4), 415-445. doi:10.1080/13506285.2013.800621

Sui, J., & Humphreys, G. W. (2014). Super-size me: self biases increase to larger stimuli. *Psychonomic Bulletin & Review*. doi:10.3758/s13423-014-0690-6

Sui, J., & Humphreys, G. W. (2015a). The Integrative Self: How Self-Reference Integrates Perception and Memory. *Trends in Cognitive Sciences, 19*(11), 719-728 doi:10.1016/j.tics.2015.08.015

Sui, J., & Humphreys, G. W. (2015b). The interaction between self-bias and reward: Evidence for common and distinct processes. *Q J Exp Psychol, 68*(10), 1952-1964. doi:10.1080/17470218.2015.1023207

Sui, J., & Humphreys, G. W. (2016). The ubiquitous self: what the properties of self-bias tell us about the self. *Ann N Y Acad Sci*, n/a-n/a. doi:10.1111/nyas.13197

Sui, J., Liu, M., Mevorach, C., & Humphreys, G. W. (2015). The Salient Self: The Left Intraparietal Sulcus Responds to Social as Well as Perceptual-Salience After Self-Association. *Cereb Cortex, 25*(4), 1060-1068. doi:10.1093/cercor/bht302

Sui, J., Ohrling, E., & Humphreys, G. W. (2016). Negative mood disrupts self- and reward-biases in perceptual matching. *The Quarterly Journal of Experimental Psychology, 69*(7), 1438-1448. doi:10.1080/17470218.2015.1122069

Sui, J., Rotshtein, P., & Humphreys, G. W. (2013). Coupling social attention to the self forms a network for personal significance. *Proc Natl Acad Sci USA, 110*(19), 7607-7612. doi:10.1073/pnas.1221862110

Sui, J., Sun, Y., Peng, K., & Humphreys, G. (2014). The automatic and the expected self: separating self- and familiarity biases effects by manipulating stimulus probability. *Attention, Perception, & Psychophysics*, 1-9. doi:10.3758/s13414-014-0631-5

Sui, J., Yankouskaya, A., & Humphreys, G. W. (2015). Super-Capacity Me! Super-Capacity and Violations of Race Independence for Self- but Not for Reward-Associated Stimuli. *J Exp Psychol Hum Percept*

Perform, 41(2), 441-452. doi:10.1037/a0038288

Swann, W. B., & Bosson, J. K. (2010). Self and identity. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (5 ed., pp. 589-628). Hoboken, New Jersey: John Wiley & Sons, Inc.

Szucs, D., & Ioannidis, J. P. (2017). Empirical assessment of published effect sizes and power in the recent cognitive neuroscience and psychology literature. *PLoS Biol, 15*(3), e2000797. doi:10.1371/journal.pbio.2000797

Tamir, D. I., & Mitchell, J. P. (2013). Anchoring and adjustment during social inferences. *Journal of Experimental Psychology: General, 142*(1), 151-162. doi:10.1037/a0028232

Tappin, B. M., & McKay, R. T. (2016). The Illusion of Moral Superiority. *Social Psychological and Personality Science*. doi:10.1177/1948550616673878

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103*(2), 193-210. doi:10.1037/0033-2909.103.2.193

Todorov, A. (2008). Evaluating Faces on Trustworthiness. *Annals of the New York Academy of Sciences, 1124*(1), 208-224. doi:10.1196/annals.1440.012

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of Competence from Faces Predict Election Outcomes. *Science, 308*(5728), 1623-1626. doi:10.1126/science.1110589

Todorov, A., Mende-Siedlecki, P., & Dotsch, R. (2013). Social judgments from faces. *Current Opinion in Neurobiology, 23*(3), 373-380. doi:10.1016/j.conb.2012.12.010

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social Attributions from Faces: Determinants, Consequences, Accuracy, and Functional Significance. *Annual Review of Psychology, 66*(1), 519-545. doi:10.1146/annurev-psych-113011-143831

Tse, C.-S., & Altarriba, J. (2012). The effects of first- and second-language proficiency on conflict resolution and goal maintenance in bilinguals: Evidence from reaction time distributional analyses in a Stroop task. *Bilingualism: Language and Cognition, 15*(3), 663-676. doi:10.1017/S1366728912000077

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *NeuroImage, 16*(3), 765-780. doi:10.1006/nimg.2002.1131

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Hum Brain Mapp, 33*(1), 1-13. doi:10.1002/hbm.21186

Turner, B. M., Rodriguez, C. A., Norcia, T. M., McClure, S. M., & Steyvers, M. (2016). Why more is better: Simultaneous modeling of EEG, fMRI, and behavioral data. *NeuroImage, 128*, 96-115. doi:10.1016/j.neuroimage.2015.12.030

Ueno, T., Fastrich, G. M., & Murayama, K. (2016). Meta-Analysis to Integrate Effect Sizes Within an Article: Possible Misuse and Type I Error Inflation. *Journal of Experimental Psychology: General, 145*(5), 643-654. doi:10.1037/xge0000159

Van Lange, P. A. M., & Sedikides, C. (1998). Being more honest but not necessarily more intelligent than others: generality and explanations for the Muhammad Ali effect. *European Journal of Social Psychology, 28*(4), 675-680. doi:10.1002/(SICI)1099-0992(199807/08)28:4<675::AID-EJSP883>3.0.CO;2-5

Van Zandt, T. (2000). How to fit a response time distribution. *Psychonomic Bulletin & Review, 7*(3), 424-465. doi:10.3758/bf03214357

Van Zandt, T., Colonius, H., & Proctor, R. W. (2000). A comparison of two response time models applied to perceptual matching. *Psychonomic Bulletin & Review, 7*(2), 208-256. doi:10.3758/bf03212980

Vandekerckhove, J., & Tuerlinckx, F. (2008). Diffusion model analysis with MATLAB: A DMAT primer. *Behavior Research Methods, 40*(1), 61-72. doi:10.3758/brm.40.1.61

Vandekerckhove, J., Tuerlinckx, F., & Lee, M. D. (2011). Hierarchical diffusion models for two-choice response times. *Psychological Methods, 16*(1), 44-62. doi:10.1037/a0021765

Vickers, D., & Lee, M. D. (1998). Dynamic Models of Simple Judgments: I. Properties of a Self-Regulating Accumulator Module. *Nonlinear Dynamics, Psychology, and Life Sciences, 2*(3), 169-194. doi:10.1023/a:1022371901259

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1-48.

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion Models in Experimental Psychology. *Experimental Psychology, 60*(6), 385-402. doi:10.1027/1618-3169/a000218

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32*(7), 1206-1220. doi:10.3758/bf03196893

Voss, A., & Voss, J. (2008). A fast numerical algorithm for the estimation of diffusion model parameters. *Journal of Mathematical Psychology, 52*(1), 1-9. doi:10.1016/j.jmp.2007.09.005

Voss, A., Voss, J., & Lerche, V. (2015). Assessing Cognitive Processes with Diffusion Model Analyses: A Tutorial based on fast-dm-30. *Frontiers in Psychology, 6*. doi:10.3389/fpsyg.2015.00336

Wabersich, D., & Vandekerckhove, J. (2014). Extending JAGS: A tutorial on adding custom distributions to JAGS (with a diffusion model example). *Behavior Research Methods, 46*(1), 15-28. doi:10.3758/s13428-013-0369-3

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology, 21*(5), 641-671. doi:10.1080/09541440802205067

Wagenmakers, E.-J., Maas, H. L. J., Dolan, C. V., & Grasman, R. P. P. P. (2008). EZ does it! Extensions of the EZ-diffusion model. *Psychonomic Bulletin & Review, 15*(6), 1229-1235. doi:10.3758/pbr.15.6.1229

Wagenmakers, E.-J., Maas, H. L. J., & Grasman, R. P. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14*(1), 3-22. doi:10.3758/bf03194023

Wager, T. D., Lindquist, M., & Kaplan, L. (2007). Meta-analysis of functional neuroimaging data: current and future directions. *Soc Cogn Affect Neurosci, 2*(2), 150-158. doi:10.1093/scan/nsm015

Wang, F., Peng, K., Bai, Y., Li, R., Zhu, Y., Sun, P., . . . Sui, J. (2016). The Dorsal Anterior Cingulate Cortex Modulates Dialectical Self-Thinking. *Frontiers in Psychology, 7*(152). doi:10.3389/fpsyg.2016.00152

Wang, F., Peng, K., Chechlacz, M., Humphreys, G. W., & Sui, J. (2017). The Neural Basis of Independence Versus Interdependence Orientations. *Psychological Science*. doi:10.1177/0956797616689079

Wegner, D. M., & Gray, K. (2016). *The Mind Club: Who Thinks, What Feels, and Why It Matters*: Viking.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41*(1), 67-85. doi:10.1016/0001-6918(77)90012-9

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical Bayesian estimation of the Drift-Diffusion Model in Python. *Frontiers in Neuroinformatics, 7*. doi:10.3389/fninf.2013.00014

Willis, J., & Todorov, A. (2006). First Impressions: Making Up Your Mind After a 100-Ms Exposure to a

Face. *Psychological Science, 17*(7), 592-598. doi:10.1111/j.1467-9280.2006.01750.x

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology, 47*(2), 237-252. doi:10.1037/0022-3514.47.2.237

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the Dominance of Moral Categories in Impression Formation. *Personality and Social Psychology Bulletin, 24*(12), 1251-1263. doi:10.1177/01461672982412001

Wrangham, R. W., & Peterson, D. (1996). *Demonic males: Apes and the origins of human violence*. New York: Houghton Mifflin.

Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. *PLoS One, 8*(7), e68910. doi:10.1371/journal.pone.0068910

Yan, C.-G., Wang, X.-D., Zuo, X.-N., & Zang, Y.-F. (2016). DPABI: Data Processing & Analysis for (Resting-State) Brain Imaging. *Neuroinformatics, 14*(3), 339-351. doi:10.1007/s12021-016-9299-4

Yan, C.-G., & Zang, Y. (2010). DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Frontiers in Systems Neuroscience, 4*(13). doi:10.3389/fnsys.2010.00013

Yeo, B. T. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., . . . Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol, 106*(3), 1125-1165. doi:10.1152/jn.00338.2011

Zell, E., & Alicke, M. D. (2011). Age and the Better-Than-Average Effect. *Journal of Applied Social Psychology, 41*(5), 1175-1188. doi:10.1111/j.1559-1816.2011.00752.x

Zhou, B., & Krott, A. (2016). Bilingualism enhances attentional control in non-verbal conflict tasks – evidence from ex-Gaussian analyses. *Bilingualism: Language and Cognition*, 1-19. doi:10.1017/S1366728916000869

Zhu, Y., Zhang, L., Fan, J., & Han, S. (2007). Neural basis of cultural influence on self-representation. *NeuroImage, 34*(3), 1310-1316. doi:10.1016/j.neuroimage.2006.08.047

胡传鹏, 邸新, 李佳蔚, 隋洁, & 彭凯平. (2015). 神经成像数据的元分析. *心理科学进展, 23*(7), 1118－1129. doi:10.3724/sp.j.1042.2015.01118

胡传鹏, 王非, 过继成思, 宋梦迪, 隋洁, & 彭凯平. (2016). 心理学研究的可重复性问题：从危机到契机. *心理科学进展, 24*(9), 1504－1518 doi:10.3724/SP.J.1042.2016.01504

李会杰, & 陈楚侨. (2007). 注意捕获的另一扇窗户——无意视盲. *心理科学进展, 15*(4), 577-586. doi:10.3969/j.issn.1671-3710.2007.04.001

林少非. (2014). *失匹配负波加工性质的研究*. (硕士), 大连理工大学. Available from Cnki

刘明慧, 张明, & 隋洁. (2014). 自我信息对知觉选择中整体优先性的调控作用. *心理学报, 46*(3), 312-320. doi:10.3724/sp.j.1041.2014.00312

王非. (2015). *辩证思维的功能与结构神经基础：神经特质视角的探索*. (博士), 清华大学, 北京.

喻丰, 彭凯平, 韩婷婷, 柴方圆, & 柏阳. (2011). 道德困境之困境 —— 情与理的辩争. *心理科学进展, 19*(11), 1702－1712. doi:10.3724/SP.J.1042.2011.01702

佐斌, 代涛涛, 温芳芳, & 滕婷婷. (2014). 热情与能力的关系及其影响因素. *心理科学进展*(09), 1467-1474.

致 谢

衷心感谢导师彭凯平教授和巴斯大学心理学系隋洁博士对本人研究的全力支持和精心指导。你们对我的鼓励与支持，让我有底气去探索未知的课题，你们对科学的态度以及探索的精神，将让我终生受益。

感谢清华大学文化与社会认知神经科学实验室各位同窗们的热情帮助和支持！王非师兄聚合整个实验室的力量，完成了神经影像-基因的大样本研究，即便现在看来，这仍然几乎是不可能任务；也感谢你在神经成像数据处理上的帮助，让我少走了许多弯路。感谢孙洋、刘冠民、王海旭、郑雯文、蔡浏阳、邬家骅、曾光、吕子旸和过继成思等多位同门在研究过程中的付出，让我十分佩服与感激。感谢杨芊、喻丰、董蕊和冯雪对我的指导和帮助，让我成长了许多。在英国牛津大学心理学系进行的 6 个月访学期，与曹一男同学的讨论让我受益匪浅，不胜感激。

感谢我爱脑科学网（52brain.com），在这里，我获得了来自无数同行的帮助与支持。正是这个平台，让我知道自己并不是唯一一个热爱探索知识的边界同时又在未知中摸索、纠结与挣扎的人。52brain，connect me with brilliant brains!

在清华大学山野协会，让我开始真正响应“为祖国健康工作 50 年”的号召，开启了“无体育、不清华”模式，让我心存感激。更加需要感谢的，在这里我幸运地遇到了与自己情投意合的人。

感谢梦迪在过去两年中的陪伴与支持，你的拼搏、善良和机智让我非常佩服，你的陪伴让我的博士生活充满开心、阳光和希望。

感谢我的父母和兄长一直以来的支持，你们是我走上科研之路的坚强后盾。

本课题承蒙国家自然科学基金资助，本课题的完成得到大量对科研具有奉献精神的实验参与者，特此致谢。

声　明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签　名：＿＿＿＿＿＿＿日　期：＿＿＿＿＿＿

附　录

附录 A　领域自评测量表

与其他同龄人相比，你觉得自己在以下方面处在什么水平？1=非常低，12=非常高

1）能力

2）身体吸引

3）物质财富

4）社交能力

5）道德

附录 B 特质词及其所属维度评定

特质词	能力	社交	道德	无法区分	其他
外向的	3.2	88.9	0	1.6	6.3
现代的	22.2	20.6	1.6	28.6	27
自私的	1.6	12.7	81	0	4.8
健谈的	42.9	54.0	0	1.6	1.6
简朴的	6.3	14.3	42.9	14.3	22.2
沉默的	4.8	76.2	0	3.2	15.9
主动的	20.6	71.4	1.6	1.6	4.8
耐心的	36.5	39.7	14.3	3.2	6.3
开朗的	7.9	85.7	3.2	1.6	1.6
成熟的	49.2	22.2	9.5	7.9	11.1
自律的	71.4	7.9	17.5	1.6	1.6
认真的	50.8	23.8	12.7	1.6	11.1
热心的	1.6	85.7	9.5	1.6	1.6
大方的	7.9	63.5	14.3	4.8	9.5
勤奋的	58.9	24.7	5.5	4.1	6.8
有趣的	28.6	55.6	0	7.9	7.9
聪明的	88.9	4.8	3.2	0	3.2
美丽的	17.5	15.9	4.8	33.3	28.6
内向的	4.8	79.4	0	6.3	9.5
传统的	7.9	22.2	22.2	17.5	30.2
无私的	4.8	12.7	81.0	0	1.6
安静的	4.8	65.1	6.3	6.3	17.5
奢侈的	11.1	9.5	33.3	20.6	25.4
聒噪的	9.5	46	12.7	9.5	22.2
被动的	30.2	46	0	9.5	14.3
急躁的	22.2	44.4	1.6	7.9	23.8
羞涩的	6.3	74.6	1.6	4.8	12.7
幼稚的	38.1	25.4	3.2	11.1	22.2
放纵的	6.3	15.9	41.3	15.9	20.6
粗心的	69.8	9.5	7.9	3.2	9.5

冷漠的	4.8	68.3	20.6	3.2	3.2
小气的	4.8	36.5	31.7	14.3	12.7
懒惰的	38.4	24.7	17.8	5.5	13.7
无聊的	12.7	57.1	1.6	9.5	19
愚蠢的	79.4	9.5	3.2	1.6	6.3
丑陋的	12.7	12.7	11.1	33.3	30.2

注：黑体的词即为属于三个维度的词，纳入随后的因子分析。

附录 C Rosenberg 自尊量表

请您阅读下面这些句子，然后选择 1-4 中的一个数字，代表您在多大程度上同意该句子的描述：

	非常不同意	有些不同意	有些同意	非常同意
1.我感到我是一个有价值的人，至少与其他人在同一水平上	1	2	3	4
2.我感到我有许多好的品质	1	2	3	4
3.归根结底，我倾向于觉得自己是一个失败者	1	2	3	4
4.我能像大多数人一样把事情做好	1	2	3	4
5.我感到自己值得自豪的地方不多	1	2	3	4
6.我对自己持肯定态度	1	2	3	4
7.总的来说，我对自己是满意的	1	2	3	4
8.我希望我能为自己赢得更多尊重	1	2	3	4
9.我确实时常感到毫无用处	1	2	3	4
10．我时常认为自己一无是处	1	2	3	4

个人简历、在学期间发表的学术论文与研究成果

个人简历

1987 年 8 月 19 日出生于湖北省监利县。

2005 年 9 月考入湖北大学系法学系法学专业，2009 年 7 月本科毕业并获得法学学士学位。

2009 年 9 月考入湖北大学心理学系心理学专业，2012 年 7 月硕士毕业并获得教育硕士学位。

2012 年 9 月免试进入清华大学心理学系攻读社会心理学学博士至今。

发表的学术论文

[1] IJzerman, H., Čolić, M., Hennecke, M., Hong, Y., **Hu, C.-P.**, Joy-Gaba, J., . . . Lindenberg, S. (2017). Does Distance from the Equator Predict Self-Control? Lessons from the Human Penguin Project (commentary). *Behavioral and Brain Science*, 40: e86. (SCI, IF = 20.4)

[2] **胡传鹏**, 王非, 过继成思, 宋梦迪, 隋洁, 彭凯平. 2016. 心理学研究的可重复性问题：从危机到契机. *心理科学进展*, 24(9), 1504–1518. (CSSCI 收录，《人大复印资料》2017 年 01 期全文转载)

[3] **Hu, C.**, Di, X., Eickhoff, S., Zhang, M., Peng, K., Guo, H., Sui, J. 2016. Distinct and common aspects of physical and psychological self-representation in the brain: A meta-analysis of self-bias in facial and self-referential judgments. *Neuroscience and Biobehavioral Reviews*, 61: 197-207. (SCI, IF = 8.802)

[4] **胡传鹏**, 邸新, 李佳蔚, 彭凯平, 隋洁. 2015 . 神经成像数据的元分析. *心理科学进展*, 23 (7): 1118-1129. (CSSCI 收录)

[5] **Hu, C.,** & Jiang, X. 2014. An emotion regulation role of ventromedial prefrontal cortex in moral judgment. *Frontiers in Human Neuroscience*, 8. (SCI, IF = 3.634)

[6] **胡传鹏**, 彭凯平, 隋洁. 2014. 道德特征在知觉匹配中的促进和唤醒效应. 北京: *第十七届全国心理学学术会议*. （博士生论坛论文三等奖）

研究成果

参与科研项目：

[1] 国家自然科学基金面上项目（31471001）. 辩证自我的神经基础.2015/01-2018/12. （参与）

[2] 国家自然科学基金面上项目（31371017）. 获得新个人重要信息的神经基础.2014/01-2017/12.（参与）